THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[SEVENTH SERIES.]

APRIL 1934.

LXI. *The Lunar Periodicity of Earthquakes.*

By CHARLES DAVISON, *Sc.D.*, *F.G.S.**.

IN any series of after-shocks the daily numbers at first decline with great rapidity. In a few days or weeks, however, the frequency-curve becomes on the whole nearly horizontal, but it is marked by a series of fluctuations, which, as Prof. Omori's well-known curves show †, occur at somewhat regular intervals of a few days. Smoothing away minor irregularities by taking 3-day means of the daily numbers, the epochs of these fluctuations become clearly defined. For instance, in the Mino-Owari earthquake of 1891, recorded at Gifu, the maximum epochs are on Nov. 20, 29, Dec. 4, 9, 15, 25, 30, Jan. 4, 12, 20, 27, Feb. 6, 15–16, 21, 29, Mar. 8, 14, 24, 29, and so on. From Nov. 20, 1891, to May 31, 1892, inclusive, there are 27 maxima, the average length of the interval between them being 7·4 days. In the Nagoya record of the same earthquake, the interval is 7·3 days. In other earthquakes the mean intervals are nearly the same, namely, 7·2 days in the Kumamoto earthquake of 1889, the Hokkaido earthquake of 1894, the Calabrian earthquake of 1905, the Marsica earthquake of 1915, and

* Communicated by the Author.

† Japan Imp. Univ. Coll. Sci. Journ. ii. plates 4–13 (1894).

the Murchison earthquake of 1929; 7·3 days in the Calabrian earthquake of 1783, the Zenkoji earthquake of 1847, the Zante earthquake of 1893, the Assam earthquake of 1897 (first Goalpara record, July 18–Sept. 10), the Tango earthquake of 1927, and the Idu earthquake of 1930; 7·4 days in the Samos earthquake of 1873 and the Kwanto earthquake of 1923; 7·5 days in the Assam earthquake of 1897 (second Goalpara record, Sept. 22–Dec. 15), and the Messina earthquake of 1908 (Messina record); 7·6 days in the same earthquake (Catania record); 7·7 days in the Charleston earthquake of 1886 and the California earthquake of 1906; and 7·8 days in the Assam earthquake of 1897 (Maophlang record). The average duration of all the above intervals is 7·4 days. In ordinary earthquakes the intervals range from 7·1 to 7·5 days, with an average of 7·3 days; in volcanic earthquakes they range from 7·1 to 7·7 days, with an average of 7·5 days.

In the valuable 'Catalogue of Earthquakes 1918–1924,' * edited by the late Prof. H. H. Turner, the last column gives the number of observations for each day that were too few in number to fix the epicentre. Many of them are records at a single station only, but the numbers of actual earthquakes are not given. Nevertheless, the results, though of less value on this account than those obtained above, are of some interest. The mean duration of the intervals for each year range from 7·1 to 7·8 days, with an average for the whole series of 7·4 days.

Thus, the average duration of all such intervals is 7·4 days, or almost exactly one-quarter of a lunation.

In the same way the series of 7-day and 14-day means of successive daily numbers of after-shocks reveal the existence of maxima occurring at longer intervals. Taking, first, the Gifu record of the Mino-Owari aftershocks (1891), the 7-day means show maxima on Dec. 2, 11, 27, Jan. 10, 26, Feb. 14, 27, Mar. 14, 30, Apr. 11 (about), 29, May 9, 25, and June 9–10. The mean of the 13 intervals is 14·7 days. In other earthquakes the mean intervals are nearly the same, namely, 13·9 days in the Calabrian earthquake of 1905, 14·2 days in the Hokkaido earthquake of 1894 and the Messina earthquake of 1908 (Messina record), 14·3 days in the Tango earthquake of

* Brit. Ass. Rep. 1928, pp. 214–304.

1927, 14·4 days in the Kwanto earthquake of 1923, 14·5 days in the Assam earthquake of 1897 (first Goalpara record), 14·6 days in the Calabrian earthquake of 1783 and the Assam earthquake of 1897 (second Goalpara record), 14·7 days in the Kumamoto earthquake of 1889 and the Marsica earthquake of 1915, 14·8 days in the Mino-Owari earthquake of 1891 (Nagoya record), 14·9 days in the Zante earthquake of 1893, and 15·0 days in the Californian earthquake of 1906, the average of these figures being 14·6 days. In ordinary earthquakes the intervals range from 14·2 to 15·0 days, with an average of 14·6 days; in volcanic earthquakes from 14·2 to 15·5 days, with an average of 14·6 days; and, in the minor disturbances included in Turner's Catalogue, from 14·4 to 15·0 days, with an average of 14·8 days. The average duration of all the intervals is thus 14·6 or 14·7 days.

In several of the above groups of after-shocks alternate maxima of the 7·4 day series are specially prominent, and they coincide roughly with the maxima of the 14·6-day series. For instance, the Gifu series has prominent maxima on Nov. 20, Dec. 4, 15, 30, Jan. 12, 27, Feb. 15, 29, Mar. 14 and 29, the average interval between them being 14·4 days. The average intervals between prominent maxima are 14·6 days in the Hokkaido earthquake of 1894 and the Marsica earthquake of 1915, and 14·7 days in the Kwanto earthquake of 1923.

The 14-day means of the series of daily numbers reveal traces, but not always such definite traces, of a period of about double the length of the latter. The Gifu series, for example, shows maxima on Dec. 9, Jan. 3, 29, Feb. 25, Apr. 1, and May 3, the mean duration of the intervals being 29·3 days. Some of the series of after-shocks are too short to give more than one or two maxima. In the longer series the mean duration is 28·6 days in the Marsica earthquake of 1915, 28·8 days in the Mino-Owari earthquake of 1891 (Nagoya record), 29·2 days in the Zante earthquake of 1893, 29·3 days in the Messina earthquake of 1908 and the Kwanto earthquake of 1923, 29·7 days in the Calabrian earthquake of 1783, 30·3 days in the Hokkaido earthquake of 1894, and 30·5 days in the Murchison earthquake of 1929; the average interval for all these earthquakes being 29·5 days. In ordinary earthquakes the intervals range from 28·2 to 30·6 days,

with an average of 29·5 days; in volcanic earthquakes from 28·2 to 30·8 days, with an average of 29·6 days; in the minor disturbances given in Turner's Catalogue from 29·2 to 30·8 days, with an average of 29·8 days. The average duration of all the intervals is thus 29·6 days, or very nearly the mean length of a lunation.

The method used above does not give individual epochs with precision, for it depends on a single series of daily numbers, which may be disturbed by other periods or causes. When, however, the limiting epochs are far apart, the small errors to which the epochs are subject affect but slightly the mean duration of the intervals.

When the times of occurrence of separate shocks are known, a similar, but more precise, method may be used. Each lunation, beginning with new moon, is divided into 28 equal intervals, and the numbers of shocks in corresponding intervals are added together. For the 29·6-day period 14-interval means are taken of such sums. For the 14·8-day period the numbers of shocks in intervals 1 and 15, and so on, are added together, and 7-interval means are taken of the sums. For the 7·4-day period the numbers of shocks in intervals 1, 8, 15, and 22, and so on, are added together, and 3-interval means are taken of the sums *. In estimating the amplitude, the average number of shocks during each interval is taken as unity. The method, however, reduces the amplitude, and the results must be multiplied by the augmenting factors 1·573, 1·583, and 1·428, for the periods of 29·6, 14·8, and 7·4 days, respectively. The epoch is indicated by the interval in which the maximum occurs, the ends of the 28th and 14th intervals corresponding to new and full moon for the period of 29·6 days, and the end of the 14th interval for new and full moon for the period of 14·8 days. This method is used throughout the remainder of this paper, except in a few series of after-shocks and some of the shocks connected with eruptions of Japanese volcanoes.

* Owing to the small number of intervals used for the 7·4-day period, the results cannot be regarded as so accurate as those for the other periods. Taking, however, at random 50 series of seven numbers each less than 100, the series of 3-number means show three maxima in 8 series, two maxima in 27, and one maximum in 15, series. In the same number of earthquake-records considered below, there are three maxima in 1 series of means, two maxima in 1, one maximum in 31, and one maximum with a very slight irregularity in 17 series.

Ordinary Earthquakes.

The earthquakes recorded in Table I. are arranged in two groups. The first include all those of which the epicentres are given in Turner's 'Catalogue of Earthquakes for the Years 1918–24'; the second those recorded at various seismological stations *. In the first group the great earthquakes are those recorded at stations more than 80° from the origin, the minor earthquakes those recorded at less distant stations. For the earthquakes with deep foci the interval taken is 1918–27, in order to secure a larger number of records than are given during 1918–24. The depths range from ·005 to ·090, with an average of ·032, of the earth's radius.

TABLE I.

Intern. Seis. Observations (1918–24).

	No. of earth-quakes.	Period. Days.	Epoch.	Ampl.
All earthquakes	2584	29·6	14 mid.	·11
		14·8	7 end	·06
		7·4	5 mid.	·03
Great earthquakes	641	29·6	12 end	·11
		14·8	6 mid.	·09
		7·4	3 mid.	·07
Minor earthquakes	1943	29·6	14 mid.	·13
		14·8	8 mid.	·09
		7·4	6 mid.	·06
N. Hemisphere	2000	29·6	13 end	·13
		14·8	7 mid.	·06
		7·4	5 mid.	·05
S. Hemisphere..................	584	29·6	ab. 16 mid.	·06
		14·8	8 end	·14
		7·4	..	..
Earthquakes with deep foci	176	29·6	15 end	·17
		14·8	2 mid.	·16
		7·4	..	..

* J. Milne, Brit. Ass. Seis. Com. Circulars, 1901–12; E. van Everdingen, Seis. Regis. in De Bilt, 1918–29; A. Cavasino, *R. Uff. Centr. Met. Geol. Boll. Sism. Micros.* 1918–24; H. O. Wood, Univ. of Califor. Publ. Seis. Sta. Bull. 1912–20.

TABLE I. (*cont.*).

Special Observatories.

Observatory.	Interval.	No. of earthquakes.	Period. Days.	Epoch.	Ampl.
Shide	1901–12	2965	29·6	1 end	·05
			14·8	13 mid.	·03
			7·4	6 mid.	·03
De Bilt	1918–29	4868	29·6	13 end	·06
			14·8	14 mid.	·02
			7·4	5 mid.	·04
Rocca di Papa	1918–24	1933	29·6	2 end	·09
			14·8	1 mid.	·03
			7·4	4 mid.	·03
Berkeley (Cal.)	1912–20	588	29·6	1 end	·08
			14·8	11 mid.	·10
			7·4	2 mid.	·04
Madras	1912–20	792	29·6	2 end	·11
			14·8	7 mid.	·06
			7·4	5 mid.	·04
Batavia	1901–09	1031	29·6	..	..
			14·8	1 mid.	·13
			7·4	3 mid.	·11
Mauritius	1901–12	951	29·6	..	..
			14·8	14 mid.	·09
			7·4	5 mid.	·09
Cape of Good Hope ...	1901–12	765	29·6	19 end	·13
			14·8	..	..
			7·4	4 mid.	·09
Christchurch (N.Z.) ...	1902–10	638	29·6	9 end	·14
			14·8	1 mid.	·17
			7·4	7 mid.	·17

It will be noticed that the epochs in the second half of this table differ as a rule from those in the first. A large number of the movements recorded at individual observatories were, however, of local origin. For instance, of 3066 earthquakes registered at Italian observatories during nearly seven years (1918, Jan. 12–1924, Dec. 24), 1652, or 54 per cent. were recorded at one station only ;

of 1933 earthquakes registered at Rocca di Papa during the same interval, 715, or 37 per cent., were recorded at that station only. On the other hand, the earthquakes included in Turner's Catalogue were, with a few exceptions, registered at several or many stations, the average number for each earthquake during 1924 being 17.

Of all the earthquakes recorded in Turner's Catalogue, 62 per cent. originated beneath the sea; of the great earthquakes, 70 per cent., and, of the minor earthquakes, 59 per cent. The elements of the three periods for earthquakes that originated beneath the sea and under land are given in Table II.

TABLE II.

Intern. Seis. Observatories (1918–24).

	No. of earth-quakes.	Period. Days.	Epoch.	Ampl.
All earthquakes (Sea)	1593	29·6	..	..
		14·8	8 mid.	·28
		7·4	5 mid.	·07
All earthquakes (Land) ...	991	29·6	13 end	·14
		14·8	1 mid.	·09
		7·4	2 mid.	·14
Great earthquakes (Sea) ...	448	29·6	12 end	·13
		14·8	6 mid.	·11
		7·4	2 end	·09
Great earthquakes (Land)..	193	29·6	3 end	·27
		14·8	2 mid.	·14
		7·4	7 mid.	·13
Minor earthquakes (Sea) ..	1145	29·6	..	..
		14·8	8 mid.	·22
		7·4	5 mid.	·10
Minor earthquakes (Land) .	798	29·6	13 mid.	·17
		14·8	1 mid.	·08
		7·4	7 mid.	·09

The principal results given by this table are :—

(i.) In each class of earthquakes, the epoch of the 14·8-day period falls at or near the times of first and last

quarters when the origins are under the sea, and at or near the times of new and full moon when the origins are under land *.

(ii.) To this opposition of epoch is due the smaller amplitudes in the first part of Table I. as compared with those in Table II. The failure of some attempts to detect the lunar periodicity of earthquakes may also perhaps be explained in this way.

The earthquakes considered in Table III. are those of certain well-known districts †. The Italian earthquakes are macroseisms or sensible earthquakes. The Tokyo record consists of earthquakes registered in that city by the Gray-Milne seismograph ‡.

One of the most interesting regions in the above table is Japan. The results are deduced from Prof. Milne's catalogue of earthquakes during the years 1885–92 §. Two great earthquakes—the Kumamoto earthquake

* The epicentres of many of these earthquakes, though submarine, lie beneath inlets or straits or groups of islands. Separating earthquakes with epicentres within 50 or 60 miles of the coast in either direction from those beyond, the epochs, when definite, are unaltered, but the amplitudes are changed. For earthquakes under the deep sea, the amplitude of the 14·8-day period is reduced to ·05, while for those under the sea near land it is ·28. For earthquakes under the land and far from the coast, the amplitude of the 29·6-day period is increased to ·41 and that of the 14-day period to ·14.

† The record of British earthquakes (1888–1912) is disturbed by the large number of shocks felt in the neighbourhood of Menstrie during limited times. If these shocks are excluded, there remain 191 earthquakes, which give the epoch of the 29·6-day period at interval 17 end (ampl. ·44) and that of the 14·8-day period at interval 7 mid. (ampl. ·33). For the earthquakes of England and Wales only (83 in number), the epochs are at intervals 17 end and 9 mid.

‡ C. Davison, 'History of British Earthquakes,' pp. 27–34 (1924); A. de Quervain, etc., Jahresb. des Schweizerische Erdbeben-dienstes, Schweiz. Meteor. Centr. Annalen, 1914–28; G. Ingrao, *R. Uff. Centr. Met. Geof. Boll. Sism. Macros.* 1918–24; and A. Cavasino, *Ital. Soc. Sism. Boll.* xxvi. pp. 125–138 (1926); xxvii. pp. 25–40 (1927); and *Boll. Sism. Macros.* 1927; *Boll. Sism. Micros.* 1918–24; D. Eginitis, *Athènes Obs. Nat. Ann.* ii. pp. 189–348 (1900); iii. pp. 336–375 (1901); iv. pp. 489–577 (1906); v. pp. 561–580 (1910); J. Milne, Japan Seis. Journ. iv. pp. 1–363 (1895); Brit. Ass. Rep. 1891, pp. 123–124; 1892, pp. 93–95; 1893, pp. 214–215; 1895, pp. 114–115; 1897, pp. 133–137; 1898, pp. 189–191; 1899, pp. 181–191; Brit. Ass. Seis. Com. Circs. i. pp. 29–30, 90–92, 142–144, 223–225; M. Sadorra Masó, 'La Seismologia en Filipinos,' Manila, 1895, pp. 75–75; U.S. Geol. Surv. Bull. nos. 68, 95, 112; F. de Montessus de Ballore, Chile Serv. Sismol. Bol. 1906–1908.

§ Reference should be made to two valuable memoirs dealing with the lunar periodicity of Japanese earthquakes, by C. G. Knott in Roy. Soc. Proc. lx. pp. 457–466 (1897), and by A. Imamura in Imp. Earthq. Inv. Com. Publ. no. 18, pp. 41–71 (1904).

of 1889, July 28, and the Mino-Owari earthquake of 1891, Oct. 28—were followed by many after-shocks, and,

TABLE III.

Regional Earthquakes.

Region.	Interval.	No. of earthquakes.	Period. Days.	Epoch.	Ampl.
Switzerland	1914–28	665	29·6	26 end	·17
			14·8	13 mid.	·19
			7·4	5 mid.	·13
Italy	1918–27	1964	29·6	28 end	·09
			14·8	7 end	·05
			7·4	2 end	·03
Rocca di Papa .	1918–24	715	29·6	2 end	·11
			14·8	7 mid.	·08
			7·4	6 mid.	·09
Zante	1894–1906	2378	29·6	15 mid.	·22
			14·8	4 mid.	·14
			7·4	1 mid.	·09
Japan	1885–89	2476	29·6	..	..
	July 28		14·8	1 mid.	·11
			7·4	4 mid.	·06
Tokyo	1891–1902	1527	29·6	28 end	·17
			14·8	8 mid.	·11
			7·4	4 mid.	·12
Philippines.....	1885–89	191	29·6	4 end	·24
			14·8	9 mid.	·13
			7·4	4 mid.	·16
California	1889–92	141	29·6	27 end	·24
			14·8	8 mid.	·25
			7·4	ab. 2 mid.	·11
Atacama (Chile).	1906–08	337	29·6	26 end	·24
			14·8	14 end	·28
			7·4	2 mid.	·10

on this account, the record used ends with the earlier date.

Table IV. refers to earthquakes of various kinds in Japan. Strong earthquakes are those which disturbed

areas of more than 5000 sq. miles, weak earthquakes those which disturbed areas of less than 100 sq. miles. Earthquakes with sound were probably of small focal depth *. Submarine earthquakes originated at various

TABLE IV.

Earthquakes of Japan (Milne).

	No. of earth-quakes.	Period. Days.	Epoch.	Ampl.
Strong earthquakes	541	29·6	27 mid.	·28
		14·8	3 mid.	·25
		7·4	1 mid.	·09
Weak earthquakes	456	29·6	28 end	·11
		14·8	1 mid.	·13
		7·4	4 mid.	·19
Earthquakes with sound ..	812	29·6	5 end	·11
		14·8	1 mid.	·14
		7·4	6 mid.	·10
Submarine earthquakes ...	2439	29·6	14 end	·14
		14·8	6 end	·08
		7·4	3 mid.	·06
District 1	361	29·6	17 end	·21
		14·8	1 mid.	·16
		7·4	5 mid.	·14
District 6	1408	29·6	27 mid.	·14
		14·8	2 mid.	·13
		7·4	5 mid.	·06
District 11	375	29·6	16 mid.	·28
		14·8	14 mid.	·17
		7·4	2 mid.	·21

distances (from 5 to 40 miles) from the coast, as a rule at about 10 miles. Milne divided the whole country into 15 districts, of which those numbered 1 to 10 lie along the east coast of the islands, and the others along the west coast. The numbers of earthquakes in the separate districts are, as a rule, too small to give

* Phil. Mag. xlix. p. 55 (1900).

satisfactory results *. Of those mentioned in Table IV., no. 1 includes the east corner of the northern island of Hokkaido, no. 6 the district around Tokyo, and no. 11 the southern region near Nagasaki. In this table the record extends from 1885, Jan. 16 to 1892, Dec. 19, except for weak earthquakes and earthquakes accompanied by sound, in which it ends on 1889, July 28, on account of the large numbers of after-shocks of the Kumamoto and Mino-Owari earthquakes.

TABLE V.

Earthquakes of Japan (Omori).

	No. of earth-quakes.	Period. Days.	Epoch.	Ampl.
Origins under land or inlets.	296	29·6	..	..
		14·8	14 mid.	·13
		7·4	5 mid.	·10
Origins under sea	287	29·6	...	..
		14·8	8 mid.	·26
		7·4	4 end	·13
Mt. Tsukuba	470	29·6	..	..
		14·8	2 mid.	·13
		7·4	4 mid.	·14

Table V. is based on Omori's catalogue of the stronger Japanese earthquakes of the years 1902–07, and his list

* The chief effect of the small number of earthquakes is to render the epoch of the 29·6-day period indeterminate. In a few districts the epochs of the shorter periods are definite, as follows :—

District.	No. of Earth-quakes.	Period. 14·8 days.	Period. 7·4 days.
2	111	1 mid.	7 mid.
3	81	12 mid.	4 mid.
4	69	..	7 mid.
5	255	4 mid.	4 mid.
9	139	14 end	3 mid.

of local shocks recorded on Mount Tsukuba in 1905 (Jan. 5–Dec. 26) *.

After-Shocks.

The lunar periods of after-shocks, as regards duration, have been considered in the introduction to this paper. Table VI. contains the elements of the periods deduced from catalogues in which the times of the after-shocks are given †.

For the Kumamoto earthquake of 1889, July 28, and the Mino-Owari earthquakes of 1891, Oct. 28, we possess tables of the daily numbers of after-shocks. The first method applied to them shows that, in both earthquakes, the maxima of the 29·6-day period fall close to the time of new moon and those of the 14·8-day period to the times of new and full moon. The same method applied to Pignataro's list of after-shocks of the Calabrian earthquake of 1783 Feb. 5, gives maxima for both periods at the same times ‡; and to both Goalpara records of the Assam earthquake of 1897 gives maxima of the 14·8-day period at the times of new and full moon.

Volcanic Earthquakes.

In considering volcanic earthquakes, the chief difficulty arises from their tendency to occur in brief, but very intense, clusters. As analysis then becomes almost useless, the present section is confined to the Hawaiian records for the years 1921, 1924, 1925, 1927, and 1928, and the Asama-yama record for 1911. The second

* Imp. Earthq. Inv. Com. Bull. ii. pp. 58–88 (1908); do., Publ. no. 22 A, pp. 1–39 (1908).

† T. Taramelli and G. Mercalli, *Roma R. Acc. Linc. Mem.* iii. pp. 88–96 (1886); *R. Uff. Centr. Met. Geod. Ann.* viii. pp. 260–289 (1888); D. Eginitis, *Athènes Obs. Nat. Ann.* ii. pp. 201–207 (1900); F. Omori, Imp. Earthq. Inv. Com. Publ. no. 7, pp. 33–51 (1902); C. S. Middlemiss, India Geol. Surv. Mem. xxxii. pp. 370–409 (1905); A. C. Lawson, 'The California Earthquake of April 18,' i. pp. 423–433 (1908); G. Martinelli, *Ital. Soc. Sism. Boll.* xvi. pt. 2, pp. 1–625 (1909); A. Cavasino, *ibid.* xix. pp. 219–291 (1915); T. Yasuda, Imp. Earthq. Inv. Com. Publ. (Jap.), no. 100 A, pp. 261–310 (1925); N. Nasu, Earthq. Res. Inst. Bull. vi. pp. 300–330 (1929); C. E. Adams, Wellington Dom. Obs. Bull. E 21, E 22.

‡ G. Vivenzio, 'Istoria de' Tremuoti avvenuti nella Provincia della Calabria ulteriore . . . . nell' anno 1783,' vol. ii. pp. i–cii. (1788). It is interesting to notice that, in the first column of his table, Pignataro inserts the dates of the four lunar phases and of perigee and apogee, though without, as he says, wishing to imply that the moon has any influence on the occurrence of earthquakes.

TABLE VI.

After-Shocks of Earthquakes.

Earthquake.	Interval.	No. of earth-quakes.	Period. Days.	Epoch.	Ampl.
Andalusia, 1884, Dec. 25	1885, Jan. 16–1886, Jan. 5	188	29·6	..	..
			14·8	7 mid.	·30
			7·4	1 mid.	·29
Riviera, 1887, Feb. 23	1887, Mar. 24–1888, Feb. 12	259	29·6	2 end	·20
			14·8	ab. 12 mid.	·22
			7·4	4 mid. or end	·09
Zante, 1893, Apr. 17	1893, June 14–Sept. 10	176	29·6	8 end	·36
			14·8	7 mid.	·17
			7·4	1 mid.	·13
Hokkaido, 1894, Mar. 22	1894, June 5–1896, Sept. 6	344	29·6	1 mid.	·14
			14·8	7 mid.	·14
			7·4	3 mid.	·13
Kangra, 1905, Apr. 4	1905, June 3–1906, Dec. 14	200	29·6	2 end	·21
			14·8	6 mid.	·23
			7·4	4 end	·21
California, 1906, Apr. 18	1906, May 23–1907, June 11	244	29·6	17 end	·22
			14·8	9 end	·16
			7·4	1 mid.	·29
Messina, 1908, Dec. 28	1909, Jan. 22–Dec. 12	135	29·6	13 end	·46
			14·8	8 mid.	·17
			7·4	..	..
Marsica, 1915, Jan. 13	1915, Feb. 14–June 17	403	29·6	14 end	·21
			14·8	5 mid.	·17
			7·4	1 end	·06
Kwanto, 1923, Sept. 1	1923, Nov. 8–1924, Jan. 6	227	29·6	2 end	·17
			14·8	14 end	·27
			7·4	5 mid.	·09
Tango, 1927, Mar. 7	1927, May 2–June 29	188	29·6	10 end	·28
			14·8	8 mid.	·23
			7·4	6 mid.	·13
Murchison, 1929, June 16	1929 July 7–Dec. 30	230	29·6	13 end	·06
			14·8	4 mid.	·14
			7·4	3 mid.	·09

method applied to these records gives the results included in Table VII.*.

By the first method applied to the Asama-yama record for 1911, it would seem that, as a rule, the maxima of the 29·6-day period fall about new moon, those of the 14·8-day period about first and last quarters, and those of the 7·4-day period about the times of the four principal phases. For the other years, 1912–16, the agreement is less close. A brief record of the Usu-san volcanic earthquakes (1918, Nov. 11–1919, Feb. 28) gives maxima

Table VII.

Volcanic Earthquakes.

Volcano.	No. of earth-quakes.	Period. Days.	Epoch.	Ampl.
Hawaii	3172	29·6	14 end	·11
		14·8	7 mid.	·11
		7·4	2 mid.	·06
Asama-yama ..	955	29·6	1 end	·39
		14·8	10 mid.	·36
		7·4	7 mid.	·27

of the 14·8-day period agreeing nearly with the times of first and last quarters.

A remarkable feature of the Usu-san eruption of 1910 was the rise of the adjoining coast. Prof. Omori gives a table showing the height on each day from Aug. 6 to Sept. 18, 1910 †. Taking 3-day means of these figures, it appears that the height oscillated, with maxima on Aug. 21, 27, Sept. 6, 12, and (about) 17, the mean duration of the intervals between them being 7·3 days. These

* Hawaiian Volc. Obs. Month. Bull. 1921–28; F. Omori, Imp. Earthq. Inv. Com. Bull. vi. pp. 115–226, 242–244 (1914; vii. pp. 217–326 (1917). In the Volcano Letter no. 170 (for Mar. 29, 1928), issued by the Hawaiian Volcano Research Association, it is noticed that, during the year 1927, " the regular recurrence of increased local seismic activity cannot be set aside as accidental. These crests of activity come at intervals of about two weeks, near the times of the first and last phases of the moon."

† Imp. Earthq. Inv. Com. Bull. vi. p. 26 (1911).

dates coincide nearly with those of the phases of the moon, namely, Aug. 20 (full moon), 27, Sept. 3, 11, and 18. The series of 7-day means shows maxima on Aug. 27–28, and (about) Sept. 8, the epochs of last and first quarters being Aug. 27 and Sept. 11.

Amplitudes of the Lunar Periods.

Though individual amplitudes cannot be regarded as of great accuracy, the average amplitudes of different classes of earthquakes, as given in Table VIII., may be of some interest.

Table VIII.

Average Amplitudes of Lunar Periods.

	29·6 days.	14·8 days.	7·4 days.
Seismological observatories ...	·12	·09	·08
Regional earthquakes	·18	·15	·11
After-shocks	·23	·20	·15
Volcanic earthquakes	·25	·23	·16

That is, the smaller the area in which the earthquakes originate, the larger the amplitude of each period.

Table IX.

Average Amplitudes of Lunar Periods.

Period. Days.	Epoch.	Aver. ampl.
29·6	New moon.	·17
	Full moon.	·18
14·8	New and full moon.	·16
	First and last quarters.	·14
7·4	Principal phases.	·18
	Midway between them.	·10

Table IX. gives the average amplitudes of the different periods corresponding to their principal epochs. It

shows that the differences are slight except for the 7·4-day period, the mean amplitude of which is nearly twice as great when the epochs fall near the principal phases as when they fall about midway between them.

Conclusions.

(i.) In the records examined, there is clear evidence of the existence of three periods of mean durations 29·6, 14·8, and 7·4 days.

(ii.) In the 29·6-day period, the maximum falls, as a rule, close to the time of new or full moon, the former epoch slightly preponderating. In the 14·8-day period, the maxima usually fall either near the times of new and full moon or near those of first and last quarters, the former being slightly in excess. In the 7·4-day period, the maxima fall in all the seven intervals, but most frequently about halfway between the times of the four lunar phases.

(iii.) It is worthy of notice that, when the maxima of the 29·6-day period fall about new moon, those of the 14·8-day period usually fall about new and full moon; while, when the maxima of the 29·6-day period fall about full moon, those of the 14·8-day period fall close to first and last quarters.

(iv.) These relations with the phases of the moon depend no doubt on the mode of origin of the earthquakes. If the earthquakes are due to movements connected with the elevation of the crust, we should expect the maxima of the 29·6-day period to coincide with new moon and those of the 14·8-day period with new and full moon. If they are caused by the depression of the crust, the maxima of the 29·6-day period should coincide with full moon, and those of the 14·8-day period with first and last quarters. The after-shocks of earthquakes are due to movements in both directions, but it would seem to be of no little significance that the maxima of the 14·8-day period should coincide with the times of first and last quarters in volcanic earthquakes, in several series of after-shocks, in the earthquakes of most maritime regions, and in the submarine earthquakes of Japan and of the whole world.

LXII. *Ferromagnetism in the Oxide obtained by Dehydration of Gamma Ferric Oxide Hydrate.* By

LARS A. WELO *and* OSKAR BAUDISCH, *Sloane Physics and Sterling Chemistry Laboratories, Yale University* *.

STUDIES † in recent years have established that the mono-hydrate, $Fe_2O_3 . H_2O$, is the only definite combination between ferric oxide and water, and that it occurs in two crystal modifications. Either modification can now be prepared artificially. The crystal transformations that take place on heating and the contrast in magnetic behaviour as the two hydrates lose water may be represented in the following scheme ‡§:

$\gamma Fe_2O_3 . H_2O$ orthorhombic ($a : b : c = 0·308 : 1 : 0·246$) paramagnetic	$\alpha Fe_2O_3 . H_2O$ orthorhombic ($a : b : c = 0·459 : 1 : 0·303$) paramagnetic
↓	↓
γFe_2O_3 cubic ferromagnetic	
↓	
αFe_2O_3 rhombohedral paramagnetic	αFe_2O_3 rhombohedral paramagnetic

From a magnetic point of view the most striking feature of these transformations is that the gamma hydrate yields

* Communicated by the Authors.

† E. Posnjak and H. E. Merwin, Am. Journ. Sci. [4] xlvii. p. 311 (1919); J. Böhm, *Zeit. f. anorg. u. allgem. Chem.* cxlix. p. 203 (1925); W. H. Albrecht, *Ber. d. Deut. Chem. Ges.* lxii. p. 1475 (1929); O. Baudisch, 'Science,' lxxvii. p. 317 (1933).

‡ The axial ratios are due to J. Böhm, *Zeit. f. Krist.* lxviii. p. 567 (1928).

§ This nomenclature is the rational one proposed by F. Haber, *Naturwissenschaften*, xiii. p. 1007 (1925), and is analogous to that in use for the isomorphous hydrates and oxides of aluminium. $\alpha Fe_2O_3 . H_2O$ corresponds to goethite and $\gamma Fe_2O_3 . H_2O$ to lepidocrocite in the modern use of these mineralogical names. Historically it would be more correct to call $\gamma Fe_2O_3 . H_2O$ goethite, since this name was first applied to the hydrate which becomes strongly magnetic on dehydration. αFe_2O_3 corresponds to hematite. γFe_2O_3 has as yet no mineralogical name. It is usually called ferromagnetic ferric oxide.

a ferromagnetic intermediate product. This interesting behaviour was recorded nearly a hundred years ago by von Kobell * and had been observed again by Plücker † in 1848. It had been forgotten in 1925 when rediscovered by Sosman and Posnjak ‡. It is with the gradual development of this product by heating and the changes in its magnetic properties that this paper will be chiefly concerned.

Preparation of Gamma Ferric Oxide Hydrate.

Most methods for preparing hydrated ferric oxide yield the alpha form. In 1928 Böhm § recognized and casually mentioned that the hydrate obtained by the action of oxygen on Fe_2S_3 is the gamma form. A year later Albrecht || showed that oxidation of a ferrous salt in dilute solution with an iodate in the presence of sodium thiosulphate also yields the gamma modification. Several newer methods have been described by Baudisch and Albrecht ¶. The material used in this study was prepared by the "pyridine method" described in the latter paper.

Pure iron powder made from iron penta-carbonyl was dissolved in concentrated hydrochloric acid until all of the acid was consumed. Iron in excess insured the absence of iron in the ferric form. The free iron was removed by filtration, and the ferrous chloride was diluted to a specific gravity of 1·019. While filtering and diluting, it was advisable to work rapidly to minimize oxidation previous to the addition of pyridine. Merck's pyridine (purity seemed to be very essential) was then added and oxygen from a tank was immediately allowed to bubble through the solution. The volume of pyridine used was about 1/30 of the volume of diluted ferrous chloride solution. The flow of oxygen was stopped in half an hour. After allowing the precipitate to settle for two minutes the upper part of the solution was syphoned off and the

* F. von Kobell, 'Grundzüge der Mineralogie,' p. 304 (Nürnberg, 1838).

† J. Plücker, *Pogg. Ann.* lxxiv. p. 321 (1848).

‡ R. B. Sosman and E. Posnjak, Journ. Wash. Acad. Sci. xv. p. 329 (1925).

§ J. Böhm, *Zeit. f. Krist.* lxviii. p 567 (1928).

|| W. H. Albrecht, *Ber. d. Deut. Chem. Ges.* lxii. p. 1475 (1929).

¶ O. Baudisch and W. H. Albrecht, Journ. Am. Chem. Soc. liv. p. 943 (1932).

remainder filtered. The precipitate was washed on the filter until chlorine free. It was then spread on paper and dried at room-temperature. The yield was found to vary from 30 to 50 per cent.

The water-content of the particular sample of gamma hydrate that was used as starting material in the present studies was 16·3 per cent. of dry weight. A portion that had been dried over calcium chloride for five months contained 15·6 per cent. of water. The magnetic susceptibility was not measured. Albrecht *, from his measurements on several samples, considers that the susceptibility of gamma hydrate is the same as that of alpha hydrate, $\chi=+42\times10^{-6}$ at room-temperature.

Dehydration and Transformation Temperatures and General Stability Relations.

Believing that our sample of gamma ferric oxide hydrate was an exceptionally pure one, and knowing that it was free from electrolytic impurities, which raise the transformation temperatures, we first obtained general information as to the temperatures at which dehydration and the transformation occur. To do this we heated the material at successively increasing temperatures for constant periods of time until nearly all was converted to alpha oxide, as indicated by feeble ferromagnetism. The pressure was atmospheric, since the electric furnace was covered with a loosely fitting lid. Data for a permeability curve were obtained after each heating. The ballistic measurements were made with an apparatus like one we have described before †. Its description and method of use will not be repeated here. But we may mention that 10 cm. was adopted as the standard length of packing instead of the 20 cm. length adopted previously. The same density of packing could not always be attained, but it never departed greatly from an average value of 0·31 g./cm.^{3}. The observed permeability values were adjusted to this packing density by the formula

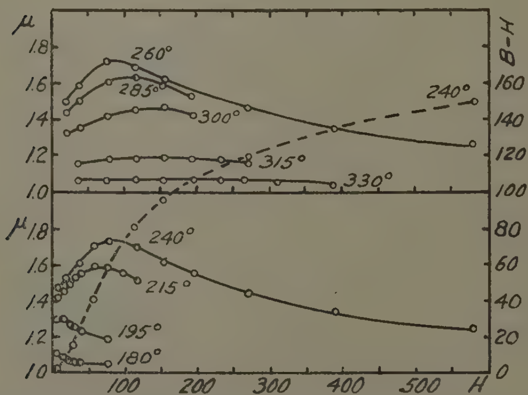
$$\mu=(\mu_0-1)\frac{0{\cdot}31}{\delta}+1,$$

* W. H. Albrecht, *Ber. d. Deut. Chem. Ges.* lxii. p. 1475 (1929).
† L. A. Welo and O. Baudisch, Phil. Mag. [6] l. p. 399 (1925).

where μ_0 is the observed value and δ is the packing density. Even then the permeabilities of this series of measurements are not strictly comparable, since the variable water-contents were not determined and allowed for.

It should be emphasized that the material was always at room-temperature when the magnetic measurements were made. Whenever, throughout this paper, there is any reference to magnetic measurements in relation to temperature, it is meant that they were made after the material had been subjected to such and such a temperature treatment.

Fig. 1.



Permeabilities after heating at increasing temperatures, and a typical magnetization curve. Heated five hours at 180° C.; at each other temperature, two hours.

The results are shown in fig. 1. The lower half refers to the dehydration stage when gamma hydrate is converted to ferromagnetic gamma oxide. The upper half shows the way in which ferromagnetism disappears as the cubic gamma oxide transforms to the rhombohedral and paramagnetic alpha oxide. A typical magnetization curve for gamma oxide fully developed from gamma hydrate by heating is also shown.

We see that a temperature of 180° C. is sufficiently high to bring about dehydration. It is seen also that

dehydration may be rapidly carried to completion at 240° C. since there was no change in permeability on heating for two hours more at 260° C. It is certain that gamma oxide is unstable at 285° C. and transforms to alpha oxide, for there is a marked decrease in permeability after heating for two hours at this temperature. The curves do not allow an inference as to the temperature at which gamma oxide becomes unstable, or rather, the temperature at which the transformation rate approaches zero. But they do show that the transformation must be slow in the vicinity of 250° C., for there was no perceptible change in permeability while heating for two hours at 260° C., unless, of course, the material passed through a state of still higher permeability.

That the transformation is slow at 250° C. was shown in a separate study of another portion of the same sample. By slow dehydration at increasing temperatures up to 250° C. an oxide was obtained with a maximum permeability of $\mu = 1{\cdot}83$. After heating at 250° C. for periods of 24, 24, and 25 hours more the permeabilities were, respectively, 1·67, 1·58, and 1·47.

At 200° C. the transformation is extremely slow. A portion of another sample of gamma hydrate was heated for 120 hours with no observable variation from the maximum permeability value, $\mu = 1{\cdot}75$. After heating for a total time of 222 hours the maximum permeability decreased to 1·70. After the total time of heating at 200° C. had been extended to 367 hours the maximum permeability was still 1·67.

Development of Ferromagnetism on Dehydration.

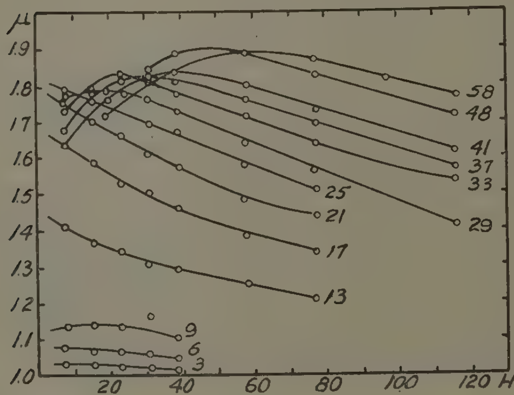
A more detailed study of the development of ferromagnetism during dehydration of gamma hydrate is shown in the permeability curves of fig. 2. Heating at 180° C. gave a convenient rate of change. The numbers in the figure refer to the total numbers of hours of heating at this temperature before taking the data for the corresponding curve. Thus, the curve marked 3 was obtained after heating three hours. The material was then replaced in the furnace and heated three hours more to give a total time of six hours and the next curve, 6, was obtained. Alternate heating and measuring were continued until no further significant changes in magnetic properties took place.

The times indicated include the time necessary for the material to reach the furnace temperature of 180° C. This usually required half an hour. In this series of measurements water determinations were made after each heating. The observed permeabilities were adjusted to the basis of pure Fe_2O_3 and to the standard density of packing, 0·31 g./cm.³, by the formula

$$\mu=(\mu_0-1)\frac{0\cdot31}{\delta}\cdot\frac{100+p}{100}+1,$$

where p is the percentage of moisture based on dry weight.

Fig. 2.



Permeabilities after heating at 180° C. for various times. Numbers indicate time in hours.

It has already been mentioned that the water-content in this gamma hydrate was reduced from 16·3 to 15·6 per cent. by drying over calcium chloride for five months. On heating at 180° C. the water-content approached the constant value of 4·4 per cent. The difference,

$$15\cdot6-4\cdot4=11\cdot2,$$

is the theoretical percentage of water in $Fe_2O_3 . H_2O$ within the experimental error; and it is inferred, from the character of the dehydration curves for hydrates

as given by Posnjak and Merwin *, that 180° C. is very close to the minimum possible dehydration temperature for this particular material.

The most interesting fact that appears from fig. 2 is that the oxide is so very soft, magnetically, when it is first formed and that it remains soft for so long a time. Up to 25 hours of heating there are no permeability maxima within the range of measurements. It is remarkable that the hardening which occurs on heating, as the curves show, should be delayed so long. The two processes, dehydration and hardening, would be expected to occur simultaneously. Instead they are separated as to time. However, this separation, in time, is not as sharp as it appears to be from a first glance at fig. 2. We estimate from the curve of water-contents that the fraction, P, of unconverted gamma hydrate after 25 hours of heating, when the material still contained 7·3 per cent. of water, was

$$P=\frac{p-4\cdot 4}{11\cdot 2}=\frac{7\cdot 3-4\cdot 4}{11\cdot 2}=\frac{2\cdot 9}{11\cdot 2}=0\cdot 26.$$

The reason that the permeability does not rise much after 25 hours is not that nearly all of the gamma hydrate is converted to oxide. The reason is that any additional permeability contributed by new oxide is compensated by a permeability decrease in the oxide formed earlier, because of the hardening. We believe that the separation as to time that did occur is due to the low temperature of dehydration. In fact, the curves of fig. 1 make this certain.

The Change in Remanence and Coercive Force.

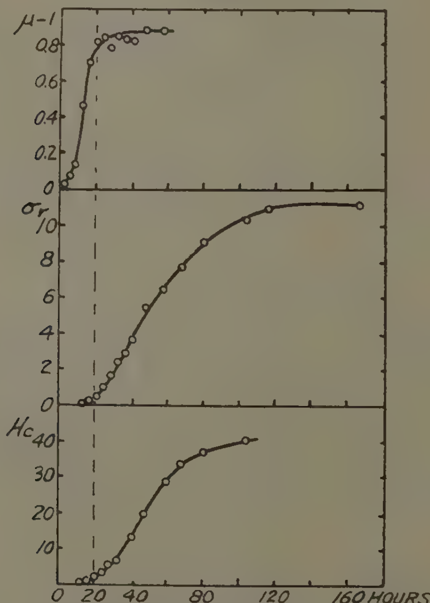
The remanences and the coercive forces were measured with the magnetometer described by Miss Wheeler †. The material, packed in a suitable container, was magnetized in the gap of a medium sized electromagnet. The field was always 1790 gausses. The container, with its magnetized material, was then mounted in the magnetometer, and the current through either the compensating coil or the magnetizing coils was gradually increased until

* E. Posnjak and H. E. Merwin, Am. Journ. Sci. [4] xlvii. p. 311 (1919).

† M. A. Wheeler, Phys. Rev. [2] xli. p. 331 (1932).

the deflected system returned to its original position. The instrumental constants given in Miss Wheeler's paper were adopted for the calculations. The original reason for the use of the magnetometer was to obtain the remanences and the coercive forces from a higher magnetizing field than was possible with the coil used for the permeability measurements. Later it was realized

Fig. 3.



Variation of $\mu-1$, remanence and coercive force with time of heating at 180° C.

that the small remanences and coercive forces encountered during the early stages of heating could not even have been detected, not to say measured, by the ballistic method.

Fig. 3 shows the variation of the remanence σ_r (specific magnetization, e.m.u.) and of the coercive force H_c as gamma hydrate is dehydrated at 180° C. It is to be remembered that the measurements were made at room-temperature after heating for the time indicated.

The upper curve shows the corresponding variation of $\mu-1$. This curve was derived from fig. 2 by a somewhat arbitrary procedure. The maximum values of $\mu-1$ were plotted whenever possible. For shorter periods of heating, when there were no maxima, values of μ, and of $\mu-1$, were obtained by extrapolating to zero fields. In spite of this arbitrariness, the facts, already mentioned in the previous section, are emphasized in the curves of fig. 3. They are that the oxide is magnetically soft when first formed and remains soft during the first 25 hours of heating, and that permeability, on the one hand, and remanence and coercive force, on the other, may be independently developed when the temperature of dehydration is low enough. Both remanence and coercive force are negligible even when the conversion of gamma hydrate to gamma oxide is three-quarters complete.

It may be explained that the curve of coercive force is only of qualitative significance because the individual determinations were not definite. It was noted that the magnetometer system continued to drift, after an increase in the external demagnetizing force, as if a condition of instability existed in the material. Demagnetization lagged behind the applied external field. Possibly, chains of magnetized particles continued to break up for some time after each change in the field. However, the curve shows the general trend of the coercive force and that coercive force changes in the same way as the remanence.

No such behaviour was noted when measuring remanence. Generally two or three minutes elapsed from the time the field of the magnet was turned off until a measurement of the remanence could be made. It is probable that higher remanence values would have been obtained if the measurements could have been made immediately. Obviously, the correct way to handle ferromagnetic powders is to fix the particles with a binder that is solid at the temperature of measurement. This is inconvenient if the material is to be heated again unless enough is available so that a new portion may be used for each treatment.

The Crystal Structures.

X-ray spectrograms were taken at various stages during the heating at 180° C. After heating for 21 hours the

lines corresponding to gamma hydrate were still strong, and it was doubtful if there were any lines corresponding to gamma oxide. After 25 hours the broad diffuse lines of gamma oxide were definitely present, but the gamma hydrate lines were still the most prominent. After 29 hours the gamma oxide lines were strong but diffuse and those of gamma hydrate could scarcely be seen. On continued heating the lines of gamma oxide became less and less diffuse. But sharp lines were not attained even after heating for 116 hours.

These observations are in complete qualitative agreement with those of Williams and Thewlis *, who studied the changes in structure as gamma hydrate is heated at a series of increasing temperatures. The crystals in their gamma hydrate, prepared in a different way from ours, had average linear dimensions of 10^{-5} cm. When crystals of gamma oxide first appeared they were extremely minute with dimensions of the order 10^{-7} cm. On continued heating they gradually grew to dimensions of the order 10^{-6} cm.

The significant information gained from our X-ray observations, when considered along with the data of Williams and Thewlis, is that the crystals of gamma oxide, when first formed, are extremely small and that they do grow even at a temperature as low as 180° C. Moreover, the time of heating required for the first appearance of gamma-oxide lines in the X-ray spectrum corresponds roughly to the time necessary for the development of any appreciable remanence and coercive force.

Discussion.

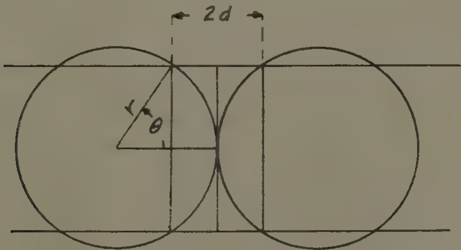
If we accept the increasing sharpness of the X-ray lines during heating as evidence that the crystals of gamma oxide coalesce to form larger ones, it is possible to account for the magnetic softness and practical absence of remanence during the early stages and for increased hardness and remanence afterwards, in terms of a variable demagnetizing factor. As to the initial softness and later hardness, the interpretation involves the application of a well-known experimental fact. It is that if the contact between ferromagnetic bodies is close enough, the

* R. D. Williams and J. Thewlis, Trans. Farad. Soc. xxvii. p. 767 (1931).

induction for the whole system of bodies approaches the value it would have if the material were continuous *.

Let us consider a chain of particles, idealized as spheres, of which only two are shown in fig. 4. r is the radius and $2d$ is the "critical" distance of magnetic contact. For simplicity we will assume that for portions of the gap exceeding $2d$ there are no mutual effects between the spheres. Within the central portion of the gap where the separation is equal to or less than $2d$ the mutual effects are the same as if there were no gap. Actually, of course, these mutual effects are continuous functions of the distance of separation. For a sphere the demagnetizing factor is $4\pi/3$. For a chain of particles, the magnetic body will be in part, in view of our assumption of perfect magnetic contact within the range $2d$, a cylinder of radius

Fig. 4.



Ferromagnetic particles in contact, considered as spheres.

$r \sin\theta$ whose demagnetizing factor will approach zero for long chains.

The ratio of cylindrical cross-section to that of the sphere will vary as r changes. Evidently,

$$d=r(1-\cos\theta), \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\cos\theta=1-\frac{d}{r}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and the cylindrical cross-sectional area

$$a=\pi(r\sin\theta)^2=\pi r^2\sin^2\theta. \quad . \quad . \quad . \quad . \quad . \quad (3)$$

When r is large compared to d, cos θ approaches 1 and

* P. Dejean, *Ann. de phys.* [9] xviii. p. 171 (1922).

$\sin^2\theta$ approaches zero, so that the cross-section of the cylinder becomes small compared to the cross-section of a sphere. On the other hand, when r decreases and approximates to d, $\cos\theta$ approaches zero, $\sin^2\theta$ approaches 1, and the cross-sectional area of the cylinder, magnetically speaking, nears that of the sphere itself.

It is readily seen that to pack a finely divided ferromagnetic powder in a tube is magnetically equivalent to filling the tube with a bundle of wires and that the ratio of the total cylindrical cross-section to the cross-section of the material as a whole is a function of the ratio d/r. In unit cross-section of packing there are $1/4r^2$ spheres and the total cross-sectional area of effectively magnetic cylinders is, from equations (2) and (3),

$$\mathrm{A}=\frac{a}{4r^2}=\frac{\pi r^2\sin^2\theta}{4r^2}=\frac{\pi}{4}(1-\cos^2\theta)=\frac{\pi d}{4r}\left(2-\frac{d}{r}\right). \quad (4)$$

When r approaches d, A, in unit cross-section, approaches $\pi/4$ and the tube is completely filled with cylinders of length equal to the packing length. For packing lengths used in our experiments the demagnetizing force would be negligible. If r is very large compared to d, A approaches zero, the magnetic cylinders become very slender as compared to the cross-section of a sphere, and the controlling demagnetizing factor becomes that of each individual sphere.

A chain need not extend the whole length of the packing. The individual chains need only to be long enough, compared to $r\sin\theta$, that the demagnetizing factor approaches zero. But chains must exist throughout all of the material.

To account for the practical absence of remanence when the particles are small and for larger values after they have grown we must make one assumption. We assume that the contact is better when an external magnetic field is applied and that the particles separate slightly when this external field is removed *. When the particles

* A method of observing and measuring such a separation is suggested by the beautiful experiment shown us by Mr. Heinz Rosenberger. A colloidal solution of gamma oxide was mounted on a slide and viewed through a microscope of 1200 diameter magnification. Straight chains were formed when a magnetic field was applied. They could be made to reach clear across the field of view. The effects of the Brownian movement were most striking. The chains would bend at some part at frequent intervals and immediately straighten themselves again.

are small the separation is large in proportion to their radii, with a proportionately large decrease in the area of magnetic contact, $a=\pi r^2 \sin^2 \theta$. Also, the field at the surface due to remanent magnetization decreases with decrease of particle size, so that the magnetic cohesion diminishes. When the particles are large the separation, as compared to r, is small, and the area of contact is not greatly affected when the magnetizing field is removed. Material in an extremely fine state of division largely demagnetizes itself because the controlling demagnetizing factor is that of the particle itself, here considered to be a sphere. When the particles are large the slender cylinders of low demagnetizing factor persist after removal of the external field.

Honda * describes an experiment which has a bearing on these contact effects. He measured the magnetization in single rows of steel balls of different sizes. In one row he used 40 balls, each 3·18 mm. in diameter, and in another he used 16 balls, each 7·94 mm. in diameter. The length of each row was 127 mm. Although the spheres were large, the magnetization was definitely higher for the smaller spheres. Honda points out in the same paper that what we here call the "effective" area of contact increases for larger intensities of magnetization. This effect is implied in our interpretation of high remanence when the particles are large. Otherwise, the cylindrical filaments would constitute too small a proportion of the whole cross-section of material. Our remanence determinations were made from relatively high applied fields, 1790 gausses.

While the contact effects that have been discussed and their dependence on particle size allow interpretation of the experimental results in terms of a variable demagnetizing factor, we are of the opinion that we are really concerned with something more fundamental to the nature of ferromagnetism. We have evidence to the effect that permeability and remanence are directly related to particle size without the intervention of the demagnetizing factor, from experiments which are free from such complications as variable perfection of lattice and variable internal strains.

In these experiments we used the highly dispersed

* K. Honda, Sci. Rep. Tôhoku, vi. p. 139 (1917).

powders that are obtained on evaporating colloidal gamma-oxide solutions to dryness. These solutions contained 50 per cent. of dextrine as carrier. The residue, therefore, consisted of gamma oxide particles embedded in and uniformly distributed throughout a much larger quantity of dextrine. The Fe_2O_3 content of the dry material was 6 per cent. The materials from two different solutions were studied. In one the oxide had been ground in a ball mill and in the other with a colloid mill. In each case the grinding was carried out wet in the presence of dextrine. Small amounts of $FeCl_3$ were used as

TABLE I.

Permeability and Susceptibility Data on Gamma Ferric Oxide (Siderac) and derived Powders for Colloids.

Substance.	Density of packing. g./cm.^{3}.	Density of oxide. g./cm.^{3}.	H, for maximum μ.	Maximum μ.	κ.	χ.
Gamma ferric oxide.	1·45	1·45	50	3·26	0·180	0·124
Same, from ball mill.	0·87	0·052	15·5	1·141	0·0112	0·215
Same, from colloid mill.	0·87	0·052	<7·7	1·136	0·0108	0·207

peptizing agents. The original gamma oxide from which these colloids had been made was a commercial product known as "siderac." Siderac is a Lefort precipitated magnetite * which has been oxidized in oxygen at 200° C. until free from ferrous iron.

Ballistic permeability and magnetometric remanence measurements were made on these "colloidal powders" and on the oxide from which they were derived, with results as shown in Tables I. and II. The volume susceptibilities were calculated by the formula $\kappa = \frac{\mu - 1}{4\pi}$.

* L. A. Welo and O. Baudisch, Phil. Mag. [7] iii. p. 396 (1927); *Bioch. Zeit.* cclviii. p. 69 (1933).

using maximum values of μ. The specific susceptibility and remanence, χ and σ_r, refer to Fe_2O_3 alone.

When considered in relation to particle size these results are exactly the same as in our study of the magnetic changes that occur as gamma hydrate is converted to gamma oxide by heating. The change in dispersion is reversed, but the direction of change in each magnetic property is also reversed. As the material is made finer by grinding, magnetic hardness decreases, permeability rises, and remanence assumes a smaller value. No contact effects are involved, since each oxide particle was embedded

TABLE II.

Specific Remanent Magnetization in Gamma Ferric Oxide (Siderac) and derived Powders for Colloids. (From H=1790 Gausses.)

Substance.	Density of packing. g./cm.^{3}.	Density of oxide. g./cm.^{3}.	σ_r. e.m.u.
Gamma ferric oxide	1·69	1·69	1·80
Same, from ball mill	0·83	0·0498	0·635
Same, from colloid mill	0·91	0·0545	1·025

in a shell of non-magnetic dextrine. Nevertheless, the susceptibilities are higher than in the undiluted oxide from which they were formed. And there is no occasion to refer to other effects that occur when particle or crystal size is changed by heating such as perfection of the lattice or release of strain. The dispersion process was purely mechanical. The decrease in hardness and especially the increase in susceptibility exclude the presence of any temperature effects that might accompany the process of grinding. A higher temperature would have increased hardness, as in our experiments with gamma oxide and gamma hydrate, and would have decreased the susceptibility because of conversion to the paramagnetic alpha form.

As to the nature of the relation between crystal or particle size and ferromagnetic properties, we can only suggest that our particles may have corresponded in size to the magnetic domains for this kind of material and that the changes in magnetic properties are due to passage through this magnetically critical size.

Summary.

Pure gamma ferric oxide hydrate corresponding to the mineral lepidocrocite, $Fe_2O_3 . H_2O$, has been prepared and studied as to the temperature-stability relations, the magnetic changes in the oxide formed by dehydration, and the changes in the X-ray diffraction pattern. Dehydration can be carried out at 180° C. At 200° C. the transformation of ferromagnetic gamma oxide to paramagnetic alpha oxide is very slow, but can be detected. The crystals of gamma oxide, when first formed, are too small to give diffraction lines, give high magnetic permeability, and have vanishingly small remanence and coercive force. On continued heating at 180° C. the diffraction lines sharpen, indicating crystal growth, and magnetic hardening occurs until remanence and coercive force reach limiting values. An interpretation in terms of a demagnetizing factor which is a function of particle size is presented and rejected. Similar magnetic results were obtained in experiments on particles of colloidal dimensions in which the variable demagnetizing factor, if operative, would have led to opposite results. It is suggested that the particles may have been passing through the magnetically critical size corresponding to the magnetic domains for this kind of material.

Acknowledgments.

We are indebted, and express our thanks, to Professor T. B. Johnson, who made this work possible, to Professor L. W. McKeehan for many valuable suggestions, and for the kind invitation to make use of a room and the facilities in the Sloane Physics Laboratory, and to Mr. R. D. Hotchkiss for preparation of the material.

New Haven, Connecticut.
September 28, 1933.

LXIII. *The Transport of Silt by a Stream.* *By* E. G. Richardson, *B.A., Ph.D., D.Sc., Armstrong College, Newcastle-upon-Tyne* *.

Introduction.

THE physics of erosion is not a subject which lends itself readily to investigation. The factors which govern the transport of sedimentary material are so many and varied that it is difficult to make laboratory experiments of a sufficiently fundamental character. There are no fewer than seven laws of similitude to consider. Consequently, it is mainly to the engineering aspects that most research work has been directed, particularly on the Continent, where the presence of strongly flowing rivers, which must be made fit for navigation by keeping their channels and estuaries free from silt, encourages investigations at nearly full scale. Very detailed investigations of this character have been made by Gilbert, Schöklitsch, and Kramer †, who have investigated the relation between the force tending to transport material from the bed ("Grenzschleppkraft"), F, and the amount of silt ("Geschiebe"), M, carried by the stream. The basis of these investigations is a semi-empirical law due to Du Boys ‡, viz.,

$$F=\rho ti,$$

where ρ=specific weight of the liquid, t=the depth, and i=the gradient of the bed of the stream. Such a law can give but a very rough representation of the dynamics of the problem, ignoring as it does several important factors. It is not surprising, then, that more or less wide departures are made from it in practice.

Of small-scale researches in the laboratory, that of Hurst § seems to be the only one dealing with the distribution of silt. Into a vessel resembling that which Joule used to measure the mechanical equivalent of heat, he

* Communicated by the Author.

† Gilbert, U.S. Geol. Survey, Prof. Paper 86 (1914); Schöklitsch, 'Schleppkraft und Geschiebebewegung,' 1914; Kramer, *Mitt. d. Preusz. Versuchsanstalt f. Wasser-und Schiff-bau*, ix. (1932). For a complete account of these and similar researches, *cf.* Wien-Harms, *Handbuch d. Physik*, IV. ii. pp. 420–447.

‡ *Ann. des Ponts et Chausées*, p. 158 (1879).

§ Roy. Soc. Proc. cxxiv. p. 196 (1929).

introduced a mixture of sand and water. With the stirring apparatus in motion, samples were run off from side taps at various heights. The concentration of sand (number per c.c. $=n$) was found to vary experimentally with height, h, according to the same law ($n=n_0e^{-\alpha h}$) as Perrin found for Brownian motion.

α depends on the size and specific gravity of the grains and the speed of stirring. It was assumed that the motion was completely turbulent.

On the theoretical side H. Jeffreys * has investigated the problem from the point of view of the upthrust which a cylindrical grain near the bottom must experience when exposed to the stream. This upthrust will vary with the size of the grain, the mean square velocity at that point, and its height above the bed. The problem is similar to the " double decker " problem in aeronautics.

Experiments.

The factors on which the erosion and subsequent transport of the bed of a stream may be expected to depend are as follows :—

(a) The velocity gradient du/dh perpendicular to the stream, particularly that on the bed itself, *i. e.*, at $h=0$;

(b) turbulence in the stream ;

(c) the grains eroded ; their size, shape, and specific gravity ;

(d) the configuration of the bed ; gradient and roughness.

In the large-scale experiments all these factors intervene in a way which does not allow of the separation of their effects. Only in Hurst's experiments are (b) and (c) the predominant factors.

It was proposed to study the transport of a sediment formed of, as nearly as possible, spherical grains of the same size and specific gravity in either laminar or turbulent flow along a channel of which the loosely compacted grains formed the bed. Thus, in a given experiment, only the first factor could vary, with the possible inter-

* Proc. Camb. Phil. Soc. xxv. p. 272 (1929). The idea of this " lift " as a factor in the transport of sediment is known to river engineers, at least, in a qualitative way. It appears to have been first mentioned by Flamant.

vention also of turbulence. The velocity gradient was to be measured by a calibrated hot-wire anemometer, and the density of distribution of the silt at different depths in the channel was to be measured in an optical manner, *i. e.*, by use of a beam of light transverse to the stream, and a photo-electric cell. When such an optical arrangement is used to find the density of *colloidal* matter a linear law connecting absorption of light and quantity of matter in suspension may be assumed *, but it is scarcely safe to assume without test that the same law will apply to diameters of grains such as those which occur in fine sand.

To obtain information on this point the experiment of Hurst was repeated. A cylindrical jar about 1 foot high, holding 3 litres of water, was filled with sieved sand and water at a concentration of 1 per cent. The average specific gravity of the sand was 2·65. Six small propellers interspersed with baffles (on one side of the jar only) stirred up the suspension in turbulent motion; a narrow beam of light traversed the other side of the vessel at various heights, and the light fell on a photoelectric cell, the illumination being measured in terms of the current through the latter. If n=the number of grains intervening, I_0=the light intensity after traversing pure water, I=the intensity after passing through the suspension

$$n = k \log I_0/I,$$

if the ordinary laws of light absorption are obeyed. When the particles are small Beer's law is obeyed, viz., k is proportional to n. Further, if Hurst's exponential law is obeyed, $\log n_0/n = \alpha h$.

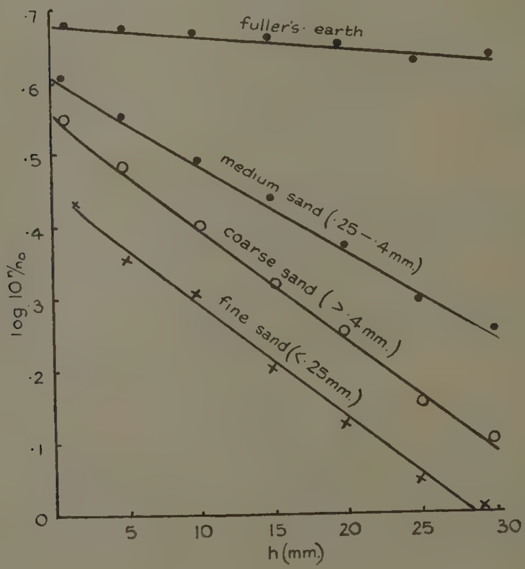
On fig. 1 the results of these experiments are plotted in the form $\log n_0/n$ against h, for different (average) grain diameters. The fact that nearly straight lines are obtained confirms the assumption that the light absorbed is, with sufficient accuracy for these experiments, proportional to n. The linearity fails at great depths, intimating departure from one or both of the above laws. That it is more likely the second than the first which fails may be judged from the reasoning developed later, pointing to abnormal concentrations near the bed. This

* *Cf.* Cheveneau and Audubert, *Ann. de Phys.* xiii. p. 134 (1920); Teorell, *Koll. Zeits.* liii. p. 322 (1930).

is confirmed by an additional test in the rotary apparatus in which a greater average concentration (2 per cent.) of sand was used. The results lie closely on the same graph as in the first instance. A decrease of specific gravity naturally diminishes α (*cf.* the lines for the fuller's earth).

The hot wire for the velocity measurements was of nickel, ·001″ diameter, mounted on a fork to allow of vertical traversing in the channel. It was calibrated by

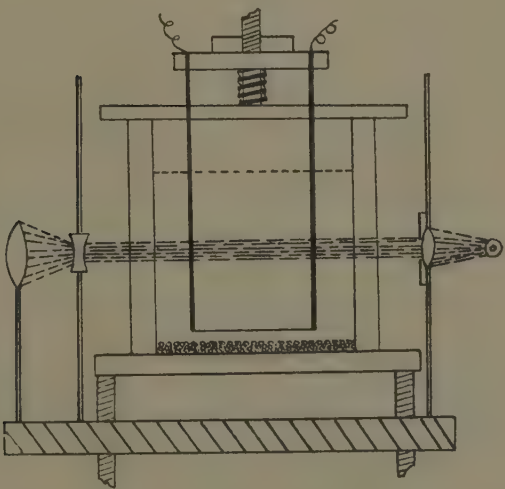
Fig. 1



being carried round a shallow trough of still water on an arm at constant speed, the resistance being measured at the same time on a simple Wheatstone bridge arrangement. The corresponding speed through the water was calculated, after a small allowance for "drift" of the water, from the speed of the whirling arm. The heating current was 0·2 amp. The calibration graph was of the familiar parabolic form. The channel was 150 cm. long, 5 cm. wide, and allowed of a depth of water up to 8 cm. It was connected at each end to large reservoirs of 25 litres

capacity. These were for the purpose of steadying the flow. The upstream reservoir was filled to a constant level from a tap, and the water was led into the channel through converging walls making a trumpet-shaped entry. From the downstream reservoir the water spilled over an adjustable notched weir. The floor of the channel was of polished wood except for a sunk section about 30 cm. long near the middle, which was filled flush with the rest of the floor with sand. The channel walls were glass plates. The measurements of velocity (u) and distribution (n) were made over a vertical section near the downstream

Fig. 2.

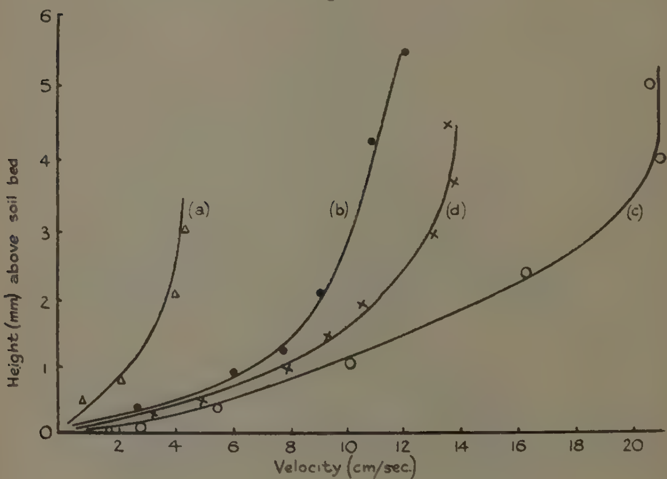


end of the model bed. The object of choosing a point about 15 diameters from the entry for the measurements was to reduce the effect which the entry might have on the flow of water through the experimental region. This is shown in section in fig. 2, with the hot wire in place, and the narrow beam of light passing across the stream into the photoelectric cell.

The first material employed for the sediment was fuller's earth. It had a specific gravity of 1·6, and consisted of particles with a mean average diameter of 0·05 mm., after allowing the larger particles to settle

in a tank and pouring off the "colloidal" matter. It was necessary to choose a light fine material of this nature if measurements of the erosion and transport under streamline conditions in the channel were to be made. Fig. 3 shows the velocity gradients in the channel at four different average speeds. At (*a*) no appreciable erosion had taken place; at (*b*) transfer of the sediment was beginning, although the motion was still streamline; at (*c*) general turbulence had intervened, and the bed was rapidly eroded. The velocity gradient in a suspension

Fig. 3.



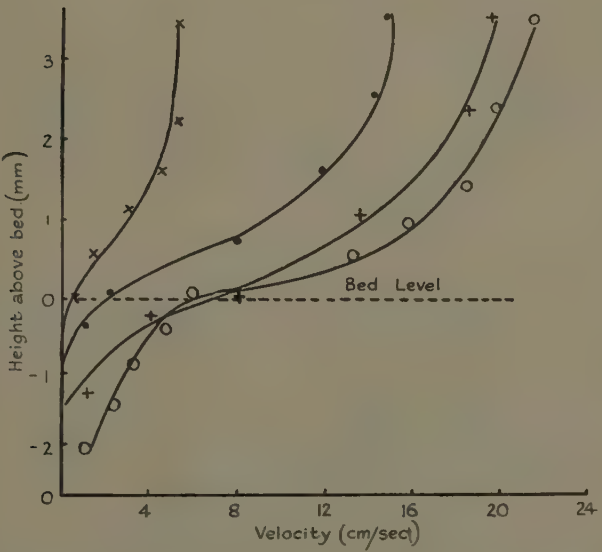
is usually steeper at the boundary than that of a homogeneous liquid * ; this can be seen on the graph in question where the line (*d*) represents the velocity gradient in water alone, at the same point in the channel at a speed intermediate between (*b*) and (*c*). This "anomalous viscosity" shown by a colloid undoubtedly makes the testing of an erosion theory difficult by introducing an abnormality into both stream-line and turbulent flow over the soil bed.

Fig. 4 gives similar data with regard to a bed of fine sand, average diameter 0·2 mm., specific gravity 2·64,

* *Cf.* Richardson and Tyler, Phys. Soc. Proc. xlv. p. 142 (1933).

where the hot-wire measurements were carried below the original bed-level to give (approximately, at any rate) the distribution of velocity in the upper part of the bed. The movement of this "saltation zone," as Gilbert calls it, has been observed by him and by Schöklitsch *. The latter used the ingenious device of letting in vertical strips of black sand among the white. The upper part of the strip was found to be inclined diagonally down stream after the flow had been taking place for a while.

Fig. 4.



Turning now to measurements in which the density of distribution was measured along with the velocity gradient, we have first a set of results (fig. 5) with a bed of average particle size ·05 mm. and specific gravity 1·60 in a stream of depth 1·5 cm. (at the section of the channel where the measurements were made). The channel floor was level and the pressure head per cm. producing the flow is given on the figure. Corresponding curves have the same mark, the velocity curves being in broken lines,

* *Loc. cit.* p. 16.

while the density curves are continuous. Fig. 6 gives similar data, but with a depth of water of 2·5 cm. Fig. 7

Fig. 5.

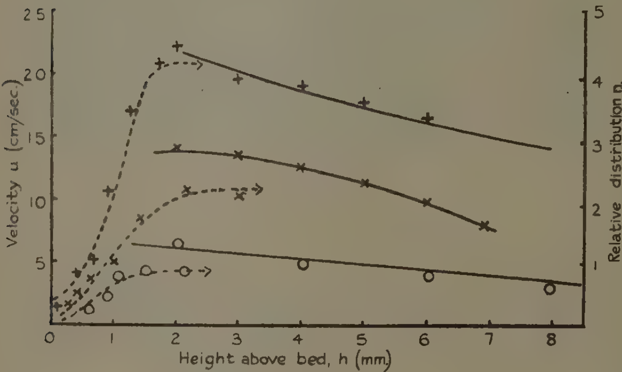
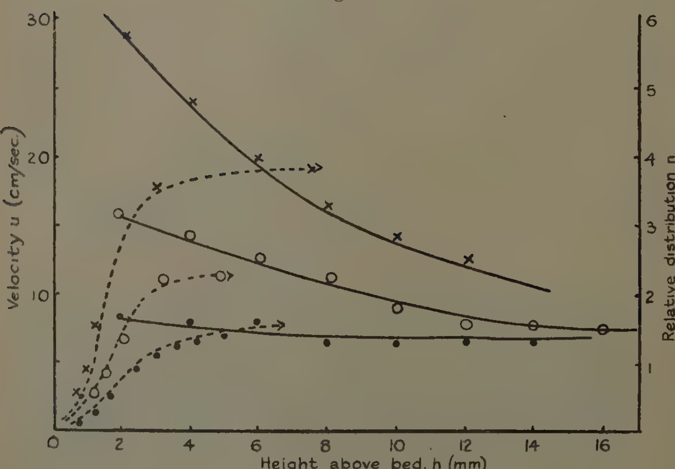


Fig. 6.

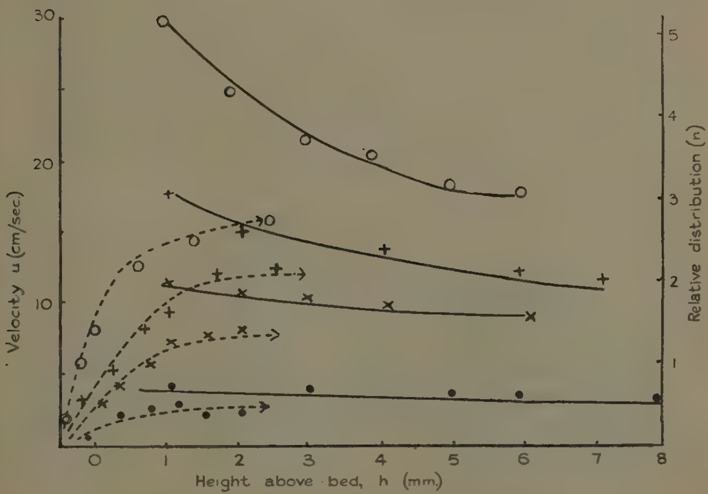


gives results with the fine sand (average diameter ·25 mm., specific gravity 2·64). In this case the flow had to be assisted by giving the bed of the channel a slope of 1 in 60

in order that considerable transport of soil should take place. The depth of the section of measurement was again 2·5 cm.

The general characteristics of the motion in the three cases illustrated are the same, viz., at the lower end of the velocity scale, where the motion is streamline, or nearly so, the density near the bed is uniform up to a certain height, after that it falls gradually as the height increases. At the other end of the scale, when the velocity

Fig. 7



is completely turbulent, the density falls in nearly exponential fashion towards the surface, resembling the distribution recorded in the "turbo-cylinder" (fig. 1). Further, of course, the total transport is continually increasing as the average velocity goes up.

Theory.

The forces on a sphere (radius a) and distant (at its centre) d from a plane wall, moving with velocity U in a stream of viscous fluid (density ρ, viscosity μ) have been

worked out by Faxen * to Oseen's approximation. The expression for the drag is

$$D=\frac{6\pi\mu a U}{1-\frac{3}{8}\frac{aU\rho}{\mu}-\frac{9}{8}\frac{a}{2d}f(x)+\left(\frac{a}{2d}\right)^3-\frac{45}{16}\left(\frac{a}{2d}\right)^4\quad 2\left(\frac{a}{2d}\right)^5}, \quad . \quad . \quad . \quad . \quad (1)$$

while the "lift" is

$$L=\frac{9}{16}\pi\rho a^2U^2\phi(x)/\left(1-\frac{27}{16}\frac{a}{d}\right):\ x=\frac{\rho U d}{2\mu} \quad . \quad . \quad (2)$$

Here

$$f(x)=1-\frac{4}{3}x+\frac{23}{16}x^2-\frac{16}{9}x^3+\frac{317}{864}x^4$$
$$+\frac{8}{9}x^5-\left\{\frac{25}{44}x^4+\ldots.\right\}\log_e(\cdot 89x),$$

and

$$\phi(x)=1-\frac{11}{8}x^2+\frac{8}{3}x^3-\frac{137}{288}x^4-\frac{8}{5}x^5$$
$$+\left\{\frac{175}{96}x^4+\ldots.\right\}\log_e(\cdot 89x).$$

These expressions hold, however, only at low values of the appropriate Reynolds number ($=x$) and are tedious to work out. It seems better to adopt the corresponding expressions relating to two spheres at a distance r apart, considering the second one as the image "in the wall" of the first. The repulsion is in this case, according to Smoluchowski †,

$$L=\frac{9}{4}\pi\mu\frac{a^2}{r}U,$$

but this again is only true for small values of Ud/μ. Oseen ‡ has given a more exact expression, *i. e.*,

$$L=\frac{9}{2}\pi\mu\frac{a^2}{xr^2}U\{1-(1+xr)e^{-xr}\}$$
$$=\frac{9}{4}\pi\mu^2U\left(\frac{a^2}{d}\right)\left\{\frac{1-(1+2xd)e^{-2xd}}{2xd}\right\},$$

* 'Thesis,' Upsala, 1912.
† Bull. Acad. Sci. Cracow, 1911.
‡ *Arkiv. f. math. astr. och. phys.* vii. no. 33 (1912).

where x has the same value as before. This reduces to Smoluchowski's expression when xd is small.

The expression in the brackets has been calculated, and shown as $f(xd)$ in the table below. In order to proceed further, and find how L varies with d, and hence to predict the density of distribution of the sediment, one must assume some law connecting U and d. If we put $U^n \propto d$, then $xd \propto Ud \propto d^{\frac{n+1}{n}}$, while $U/d \propto d^{-\frac{n-1}{n}}$. Thus, if the linear law $U \propto d$ should hold, U/d is constant and $xd \propto d^2$. We have then only to take the figures in the second column and compare them with the square roots of those in the first to find how the lift varies with distance of the particles from the bed. If $n=2$, we have to put $d \propto (xd)^{2/3}$, and divide the corresponding value of the function by $d^{1/2}$. When n is large, $xd \propto d$ itself, while $U/d \propto \frac{1}{d}$, therefore we divide the value of f by d to get the lift.

xd.	$f(xd)$.	$n=1$.		$n=2$.		$n=\infty$.	
		$d \propto (xd)^{1/2}$.	$L \propto f$.	$d \propto (xd)^{2/3}$.	$L \propto f/d^{1/3}$.	$d \propto (d)$.	$L \propto f/d$.
·2	·15	·45	·15	·34	·26	·2	1
·5	·20	·7	·20	·63	·26	·5	·4
1	·26	1	·26	1	·26	1	·26
2	·30	1·41	·30	1·6	·235	2	·15
3	·265	1·73	·265	2·1	·183	3	·13
4	·225	2	·225	2·5	·142	4	·11
6	·16	2·45	·16	3·3	·088	6	·08

The points to notice in this table are :—

1. The function has a maximum, which when $n=1$ gives a maximum lift at $d=1·41$.

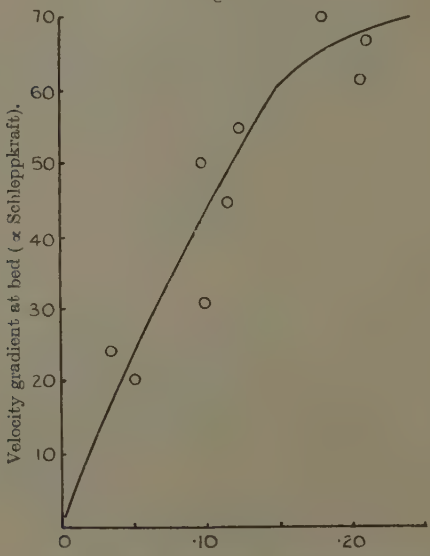
2. When $n=2$ this peak is smoothed out, and

3. When n is large the maximum has disappeared, the lift falling continuously with d, in nearly exponential fashion.

In any case, the image method fails at very low values of d, owing to the fact that it does not make the velocity zero at the wall itself.

Beside the lifting force there is a downward force $W = \frac{4}{3}\pi a^3(\rho - \rho')g$, where ρ' is the density of the liquid. The particles should find a position where these two forces are in equilibrium. It should further be noted that in the formula for L, U is the velocity relative to the fluid. As the particles are swept up from the bed they will tend to acquire the velocity of the liquid, their terminal velocity depending upon an equilibrium between the pressure

Fig. 8.

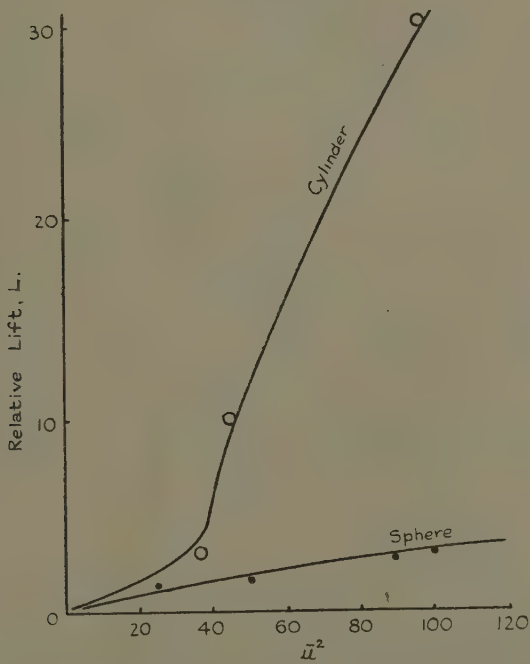


of the liquid and the drag D. In spite of these limitations we can see the data of the table reflected in the exponential results. At the lower speeds we see the flat topped density distribution characteristic of streamline flow (n small). This passes over into an exponential distribution as general turbulence is reached (n large), in which case the formulæ of Oseen fail to apply.

The onset of turbulence was observed by the Osborne Reynolds method of allowing a filament of coloured fluid to enter the stream near the mouth. When the depth

was 1·25 cm. the critical average velocity was about 6 cm./sec.; at 2 cm. it was 4·5 cm./sec., corresponding to a Reynolds number of $uh/\nu = 800$ approximately *. The velocity at which turbulence *ceases* is a little lower, a fact which influences the settling of sediment once it is raised.

Fig. 9.



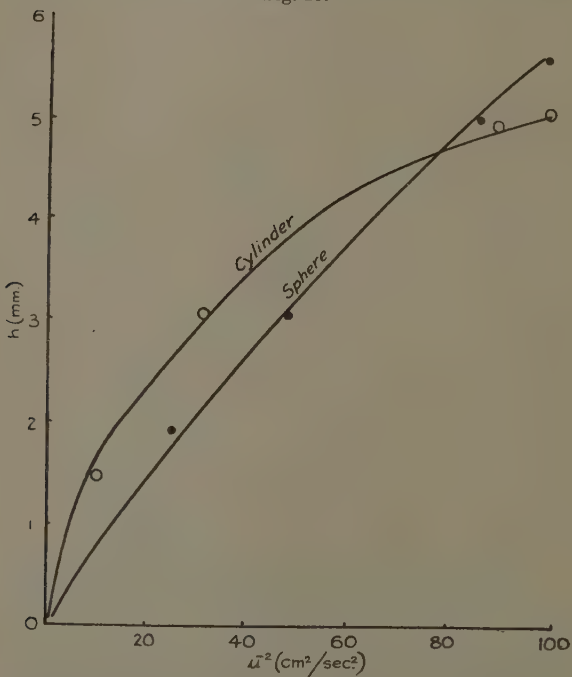
By integrating the distribution curves for a given soil it should be possible to correlate the force per unit area on the bed $\left[=\mu\rho\left(\frac{du}{dh}\right)_{h=0}\right]$, with the total amount of silt transported per unit time. This has been done for the two depths with the fuller's earth (fig. 8). One observes the marked increase in rate of transport which occurs

* *Cf.* Hopf, *Ann. d. Phys.* xxxii. p. 777 (1910).

when turbulence is reached, corresponding to a velocity gradient of 50–60 sec.$^{-1}$ at the bed.

The resistance of the small particles which we have been discussing can be expected to obey the law of the first power of the velocity, at the lower speeds at any rate. But in the case of the large pebbles which form a con-

Fig. 10.



siderable part of a river bed, eddying motion will occur behind them even at low speeds, so that the motion will be governed by u^2 rather than by u. Some experiments were made on large objects, *i. e.*, a sphere ($\frac{3}{4}''$ diameter) and a cylinder ($\frac{3}{4}''$ diameter, $2''$ long). These were arranged on the end of a force balance so that they were immersed just to touch the bottom of the channel. When the stream was set in motion, two quantities were observed:

(*a*) the height to which the body rose, (*b*) the movement of the counterpoise along the balance to keep the body in contact with the bottom. These are plotted on figs. 9 and 10 respectively. Both quantities are nearly proportional to the square of the average stream velocity, except at low speeds.

The drag on a sphere at the bottom of a channel has been measured by Schöklitsch *, who finds that this, too, is proportioned to u^{-2} within the usual range of velocities. The resulting motion of a stone on the bottom is compounded of a lift—which diminishes as the stone rises, partly because of the fall of lift with height, but also because its velocity relative to the stream diminishes—and a drag. Its resultant motion is usually a series of jumps, like a toy balloon blown along the ground. There is also a rotation due to the difference in stream velocity above and below the stone. An attempt to observe this with an axially pivoted cylinder in our model channel was unsuccessful, the torque evidently being too small.

To summarize these investigations: it has been found that the transport of silt is in the main due to turbulence in the stream, under which conditions the vertical distribution is nearly exponential, and the lifting force ("Schleppkraft") is proportional to the square of the velocity. Below the critical speed the silt is moved to a much less extent, and the movement is confined to the smaller particles in the lower strata of the stream.

LXIV. *The Resistance of Manganese Arsenide.*

By L. F. Bates, *Ph.D.*, *D.Sc.*, *Reader in Physics*, *University College, London* †.

Introduction.

It has long been known that ferromagnetic metals exhibit peculiar changes of resistance with rise in temperature, although these changes have only within recent years been subjected to adequate systematic investigation. Thus, the variation of the resistance of a nickel wire with temperature in the absence of an applied

* *Loc. cit.* p. 43.

† Communicated by the Author.

magnetic field was carefully examined by Gerlach and Schneiderhan [1]. They found that at temperatures appreciably higher than the ferromagnetic Curie point the resistance varied in the same way as that of a normal, *i. e.*, non-magnetic, metal, while below the Curie point the resistance at any given temperature was the normal resistance decreased by an amount directly proportional to the square of the spontaneous magnetization, *i. e.*, to the magnetic energy, at that temperature. The spontaneous magnetization is usually defined as the magnetization of the elementary portions of a substance due to the presence of an internal field. It may be increased by the application of a sufficiently strong external field, the resultant magnetization then being termed the true magnetization.

Ferromagnetic metals also exhibit characteristic changes of resistance when placed in magnetic fields. The variation of the resistance of wires of such metals has been examined in detail by Potter [2], Gerlach [3], and Engler [4], while Gerlach [5] has discussed the changes observed with single crystals of iron. In all cases in a strong magnetic field a decrease in resistance is observed. This decrease is due to an increase in the true magnetization of the metal produced by the field, and is directly proportional to the change in magnetic energy.

In view of the pronounced ferromagnetic properties of manganese arsenide, it was hoped to make similar examinations of the variation of its resistance with rise in temperature and the application of external fields. Unfortunately, the difficulty of producing the material in a suitable form has caused the examination to be limited to the variation of the resistance with temperature alone.

Experimental Details.

The preparation of rods of manganese arsenide has been fully described elsewhere [6], [7]. Special methods of mounting the rods for resistance measurements had to be devised, as they were short and brittle and about 2 to 4 mm. thick.

In the experiments to be described the resistivity of the material was not determined ; it was merely required to know how the resistance of a chosen length of an arsenside rod varied with temperature. Measurements

were therefore made by the well-known shunt potentiometer method of comparing low resistances. The potential falls across a standard 1000 ohm resistance and across an adjustable dial resistance connected in one circuit, were respectively balanced against those across a standard 1 ohm resistance and across a chosen length of an arsenide rod connected in a second circuit. Each arsenide rod was therefore provided with two potential and two current leads. Several modes of attaching these leads were tried, and one of the more successful will now be briefly described.

A short piece of ebonite tube, practically a ring of ebonite, whose internal diameter was approximately that of an arsenide rod, was split diametrically. Two pairs of small holes were drilled in one half of this ring. Two fine copper wires were threaded one through each pair of holes, so that each wire formed an arc of a circle on the inner surface of the ebonite.

One end of an arsenide rod was now placed between the two halves of the ebonite ring, which were then pressed together by a phosphor bronze clip. The copper wires were now in contact with the rod, and the wire more remote from the middle of the rod was used as a current lead and the other wire as a potential lead. A similar attachment was provided at the other end of the rod. This method of providing current and potential leads seems a very useful one, as for longer rods than those available in these experiments separate ebonite rings could conveniently be used for the current and potential leads.

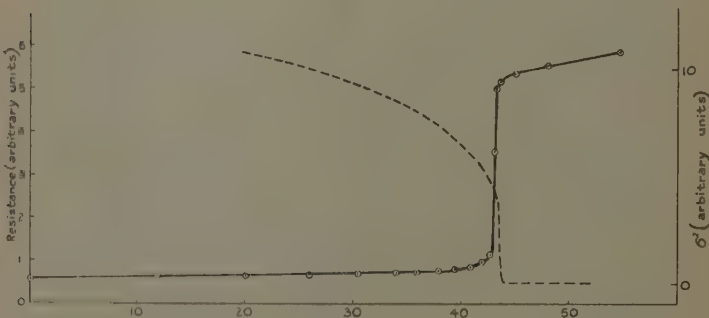
Finally, a method which was used only for the shortest specimens must be described. Two narrow cuts were made parallel to each other on the surface of a piece of thick fibre. In these cuts two fine wires were pressed so that the surfaces of the wires protruded slightly above the surface of the fibre. These wires served as the potential leads and on them was placed the arsenide rod, which was held down at each extremity by a phosphor bronze clip, screwed to the fibre, to which a current lead was attached. The method was obviously not an ideal method of attaching the leads, for it could not permit free expansion of the rod as in the previous methods, but it gave results which were sufficiently reliable in view of the large changes in resistance described below.

To make resistance measurements at various temperatures the mounted specimen was placed inside a thick copper tube which traversed a well-lagged copper tank containing oil which was electrically heated and stirred. The oil was maintained at any chosen temperature and, after a sufficient interval of time, the resistance of the specimen was measured and its temperature read on a calibrated mercury thermometer placed with the bulb practically touching the specimen. Before commencing each set of measurements the specimen was first cooled to 0° C.

Results.

A typical set of measurements obtained with a specimen, (Y 2), which had not been heated since it was cast,

Fig. 1.

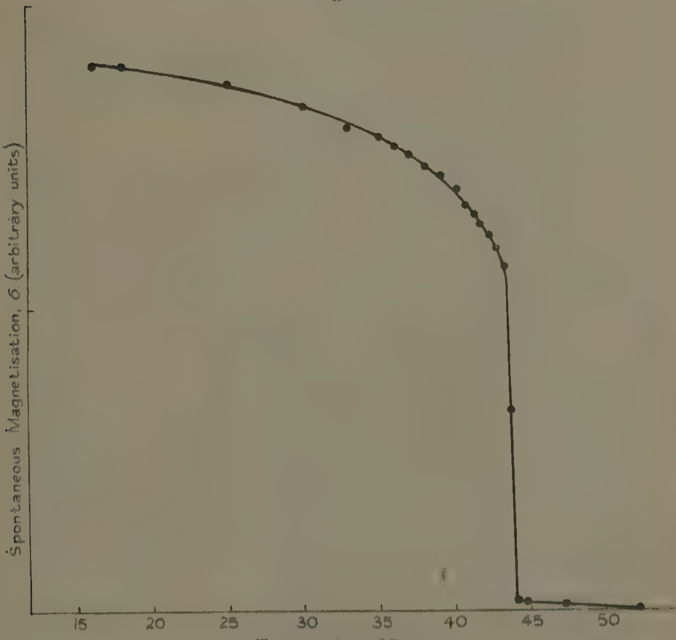


is shown by the unbroken curve of fig. 1. The curve of fig. 2 shows how the spontaneous magnetization of a similar specimen measured as described in an earlier paper, varies with temperature. These curves could not be obtained simultaneously with the same specimen; they were obtained with two specimens from the same casting.

The portion of the curve of fig. 1 which lies below 40° C. was reproducible provided that this temperature was not exceeded. In the neighbourhood of 43° C., however, a very sudden change in resistance occurred, which was followed by a more gradual change between 45° and 50°.

The slope of the last portion of the curve varied considerably from specimen to specimen. The complete curve of resistance with temperature should also be compared with the curve showing the variation of the square of the spontaneous magnetization, σ^2, with temperature, which is shown by the broken curve of fig. 1. It is clear that pronounced changes in σ^2 extend

Fig. 2.



over a much greater temperature range than the resistance changes, and there is no sign of a simple direct relation between them. This behaviour was always observed with the first sets of measurements with specimens of a given sample of arsenide, and will be further discussed below.

In any case such exact numerical agreement between the change in the magnetic energy and the change in the resistance as was obtained by Gerlach and Schneiderhan

in the case of nickel, cannot be expected in the case of the arsenide specimens. For, even when the substance is raised to temperatures above the ferromagnetic Curie point, magnetic changes still occur [8], and the substance cannot then be treated as a normal metal. Moreover, exact numerical agreement can hardly be expected, when a displacement of the Curie point by about 0·1° C. would make a very great difference to the agreement, and the magnetic and resistance curves are not obtained in the same experiment. Again, once the substance has had its temperature raised above, say, 42° C., certain permanent changes are produced, and the curves of figs. 1 and 2 cannot be reproduced when the specimen is cooled to 0° C. and heated once more.

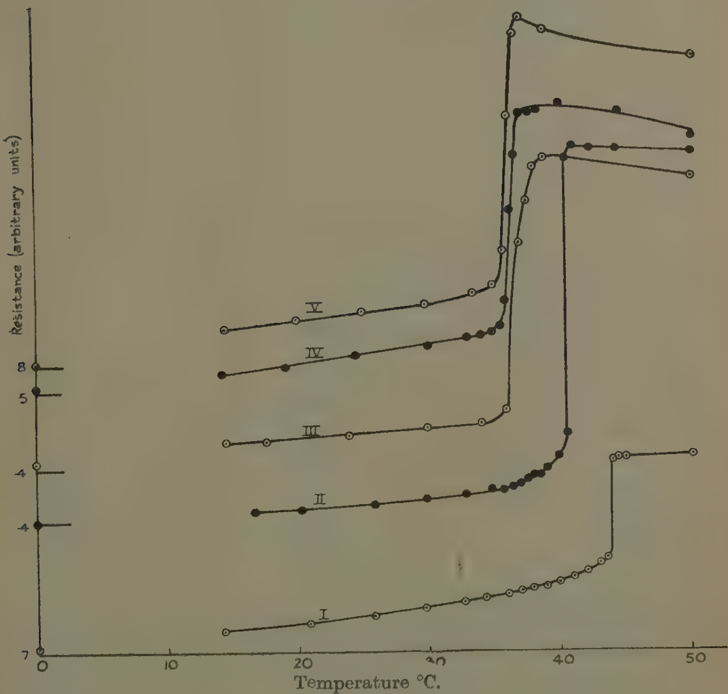
There is no doubt that at least two processes are responsible for these resistance changes, viz., that properly attributed to the changes in magnetic energy and that presumably due to the formation of very fine cracks or fissures in the specimen. Fairly large volume changes accompany the abrupt changes in ferromagnetism, and these naturally favour the formation of fissures and their enlargement during successive transitions from the paramagnetic to the ferromagnetic state.

However, a certain amount of interesting information was obtained in this investigation. Thus with a specimen of the same yield as that called B*d* 2, whose magnetic properties have been previously described [7], the set of curves shown in fig. 3 was obtained. In each case values of the resistance were obtained at temperatures from 0° to 50° C., after which the specimen was rapidly cooled to 0° C., the curves I., II., III., IV., and V. being obtained in that sequence on different days.

In fig. 3 the values of the resistance are plotted in arbitrary units ; the ordinates of the successive curves are displaced, and, in every case must be multiplied by an appropriate factor, the factors being respectively 1, 2, 6, 10, and 10. The resistance of the specimen at 0° C. increases considerably from curve to curve, and the magnitude of the resistance changes greatly increases. It will be noted that the last portions of curves IV. and V. exhibit a negative slope ; in addition, the resistance here showed a slow decrease with time. Similar sets of curves were obtained with other specimens of manganese arsenide.

The most striking feature of the curves of fig. 3 is the progressive lowering from curve to curve of the temperature at which the sudden increase in resistance takes place until the limiting condition of curve V. is attained. The magnetic and thermoelectric properties change in a corresponding manner [6], [7].

Fig. 3.



Since curve V. represents a limiting condition it has been plotted as the unbroken curve in fig. 4, where the broken curve shows the variation of the square of the spontaneous magnetization with temperature for precisely the same specimen. It is abundantly clear that the changes in resistance are now closely allied to the changes in magnetic energy, although in view of the remarks

entered above there is little point in seeking an exact numerical relationship.

Now, manganese arsenide shows a temperature hysteresis. Accordingly, in one experiment the temperature

Fig. 4.

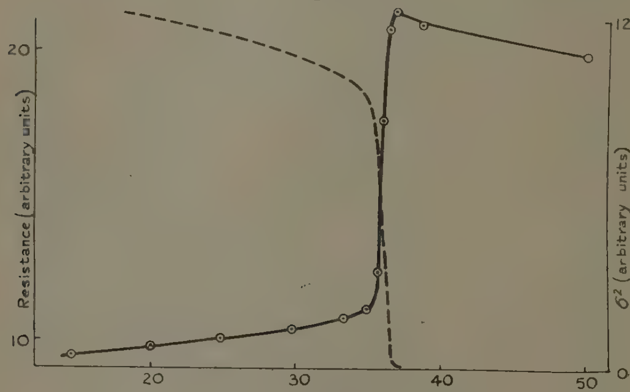
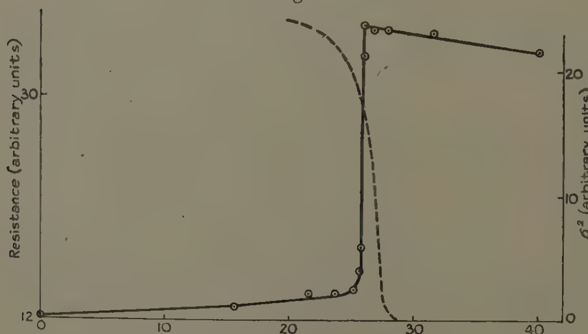


Fig. 5.

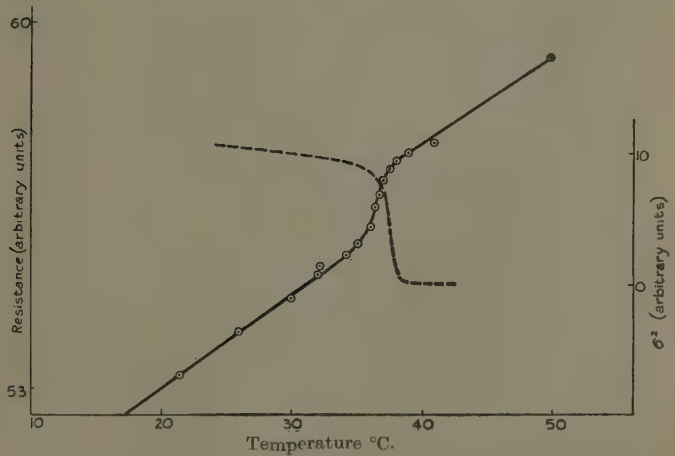


of the rod was raised fairly slowly to 50° C., where it was maintained for about half an hour. It was then allowed to cool slowly, the resistance being measured at chosen temperatures. These values are shown by the unbroken line of fig. 5, while the broken line shows the corresponding

variation of the square of the spontaneous magnetization. In this case, while there is clear evidence that the decrease in resistance closely follows the change in magnetic energy, there definitely appears to be a temperature lag of about 1° C., the ferromagnetic properties reappearing before the resistance suddenly changes.

On account of its interest in the earlier communications [(6), (7)] the resistance measurements for the specimen *Aa* are reproduced in fig. 6. Unfortunately, this specimen has been broken accidentally and only a short piece was available for the resistance measurements.

Fig. 6.



The curve of fig. 6 shows that this specimen behaves much more like nickel than any of the other specimens, which had been subjected to much higher furnace temperatures in the course of their preparation. The broken line of fig. 6 shows the variation of the square of the magnetization with temperature.

Discussion.

Evidence has been obtained which shows that the resistance of manganese arsenide, in common with that of other ferromagnetic substances, is intimately connected with its energy of spontaneous magnetization.

The relation which satisfactorily represents the results obtained with nickel may be written

$$R_T = R_{nT} - C \cdot \sigma^2,$$

where R_T is the resistance of a specimen of nickel at a temperature T, R_{nT} is the resistance of a normal, *i. e.*, non-magnetic, metal at the same temperature, and C is a constant independent of the temperature.

This relation certainly does not hold for specimens of manganese arsenide which are being demagnetized by heat for the first time, and it is concluded that in this case C varies with the temperature. The curves of fig. 4 show, however, that after the substance has been caused to pass several times from the ferromagnetic to the paramagnetic state, a limiting condition is reached in which the simple relation is much more closely satisfied.

In suggesting that the constant C varies with the temperature, it is not necessarily implied that manganese arsenide provides a case in which there is no simple direct relation between the resistance and the energy of spontaneous magnetization. For the quantity $C \cdot \sigma^2$ may be replaced by the expression $C' \cdot (n \cdot \sigma^2)$, where n is the constant of the internal field. Although it is generally postulated that n is a constant independent of the temperature, evidence from specific heat measurements indicates that n varies with temperature. If, therefore, the change in resistance is strictly proportional to the change in magnetic energy, the above experiments indicate that the constant of the internal field varies with temperature.

Summary.

The variation with temperature of the resistance of manganese arsenide has been determined, using a method suitable for measurements with specimens in the form of short brittle rods. The connexion between the observed resistance changes and changes in the energy of spontaneous magnetization is discussed.

My thanks are due to Professor E. N. da C. Andrade for the many facilities afforded me, and to the Government Grants Committee of the Royal Society for a grant with which the magnet [9] used in these investigations was constructed.

Bibliography.

(1) W. Gerlach and K. Schneiderhan, *Ann. der Phys.* vi. p. 772 (1930).
(2) H. H. Potter, Proc. Roy. Soc. A, cxxxii. p. 560 (1931); Phil. Mag. xiii. p. 233 (1932).
(3) W. Gerlach, *Ann. der Phys.* xii. p. 849 (1932).
(4) E. Englert, *Zeit. für Phys.* lxxiv. p. 747 (1932); *Ann. der Phys.* xiv. p. 489 (1932).
(5) W. Gerlach, *Ann. der Phys.* xiv. p. 613 (1932).
(6) L. F. Bates, Phil. Mag. xiii. p. 393 (1932).
(7) L. F. Bates, Phil. Mag. xvi. p. 657 (1933).
(8) L. F. Bates, Phil. Mag. viii. p. 714 (1929).
(9) L. F. Bates and B. Lloyd Evans, Proc. Phys. Soc. xlv. p. 425 (1933).

LXV. *The Deuton and Disintegration.* By HAROLD J. WALKE, *B.Sc., Mardon Research Scholar, the Washington Singer Laboratories, University College, Exeter* *.

HEISENBERG † has explained several facts concerning the existence and stability of isotopes by assuming that atomic nuclei are composed of α-particles, zero or one proton, and neutrons. An essential feature of Heisenberg's theory is that the force of attraction between a neutron and a proton (which is regarded as of the nature of a polarizing force) exceeds that between two neutrons. It is the purpose of this paper to show that this force may be more than a polarizing force, and that a neutron-proton combination may exist within the lighter nuclei as a stable nuclear sub-unit.

The nucleus of the hydrogen isotope of mass 2 consists of such a combination of a neutron and a proton, and as this unit is stable and can enter the nuclei of certain elements without itself disintegrating, as has been shown by Lewis, Livingston, and Lawrence ‡, and Oliphant, Kinsey, and Rutherford §, it is assumed that this particle is a nuclear component, and may exist within more complex nuclei. Thus the greater stability of the proton-neutron combination is to be thought of as due not to some polarizing attraction, but to the mass defect of the nucleus of heavy hydrogen. This combination, which has

* Communicated by Prof. F. H. Newman, D.Sc.

† W. Heisenberg, *Zeit. f. Phys.* lxxviii. p. 150 (1932).

‡ Lewis, Livingston, and Lawrence, Phys. Rev. xliv. p. 55 (1933).

§ Oliphant, Kinsey, and Rutherford, Proc. Roy. Soc. A, cxli. p. 722 (1933).

been called the "demi-helion," is more generally known as the deuton.

As, however, Lawrence, Livingston, and Lewis * have shown that deutons, when accelerated by high potential, break up into neutrons and protons when in collision with heavier nuclei, it is assumed that, although the combination is stable, its components are not as tightly bound as those of the α-particle, and in consequence it is a secondary nuclear component.

Landé † has supported and extended Heisenberg's model, and has developed a nuclear scheme in which he stresses more than the latter the rôle which the α-particle plays in nuclear structure. Following Landé, it is assumed in this paper that the nuclei of the lighter elements from helium to sulphur are made up of the maximum number of α-particles, zero or one proton and zero, one or two neutrons. However, Landé, from isotopic considerations, regards the proton as free, the neutrons in these nuclei being bound into a single shell which, when complete, is populated by two neutrons. In contrast I suggest that the nuclei are to be regarded as consisting of a central tightly bound core of α-particles, the less tightly bound proton and neutrons being outside the core. Where a proton and neutrons exist within the nucleus it is postulated that the neutron-proton sub-unit or deuton is formed.

The deuton will, therefore, tend to remain within the nucleus in certain cases of disintegration, but, on account of the still greater stability of the α-particle, two deutons within the nucleus will tend to unite to form an α-particle. This will occur in certain cases of bombardment by protons and deutons which may be regarded as interacting directly with these less tightly bound particles. Such α-particles will be formed within the nucleus, but outside the core, and will be ejected with the energy set free by the defect of mass emitted in the formation of the new unit. If the α-particle is not formed in this way outside the core, it is difficult to see why it does not remain in the nucleus forming the next stable α-particle configuration, the energy set free being emitted in the form of a nuclear γ-ray. However, on account of the greater penetrating

* Lawrence, Livingston, and Lewis, Phys. Rev. xliv. p. 56 (1933).
† Landé, Phys. Rev. xliii. p. 620 (1933).

power of the α-particle and the neutron, it is probable that in certain cases of disintegration these particles interact with the core.

Adopting the notation S^M_N for the elements (S=chemical symbol, M=mass number, and N=atomic number) and using the letters $\alpha=\alpha$-particle, n=neutron, p=proton, $[\alpha]=\alpha$-particle in the core, $p+n$=the deuton, we may indicate the structure of the lighter nuclei thus :

$$H^1_1=p\,;\quad H^2_1=\overline{p+n},$$

$$H^4_2=\alpha,$$

$$Li^6_3=\{[\alpha]+(\overline{p+n})\}\,;\quad Li^7_3=\{[\alpha]+(p+n+n)\}\,;$$

the { } brackets indicating the complete nucleus, the [ ] brackets representing the core, and $(\overline{p+n}+n)$ indicating that the two neutrons and the proton are regarded as distinct from the core, though one neutron and the proton $\overline{p+n}$ form a stable component. Similarly we have

$$Be^8_4=[2\alpha]\,;\quad Be^9_4=\{[2\alpha]+(n)\}.$$

$$B^{10}_5=\{[2\alpha]+(\overline{p+n})\}\,;\quad B^{11}_5=\{[2\alpha]+(p+n+n)\}.$$

$$C^{12}_6=[3\alpha]\,;\quad C^{13}_6=\{[3\alpha]+(n)\}.$$

$$N^{14}_7=\{[3\alpha]+(p+n)\}\,;\quad N^{15}_7=\{[3\alpha]+(\overline{p+n}+n)\}.$$

$$O^{16}_8=[4\alpha]\,;\quad O^{17}_8=\{[4\alpha]+(n)\}.$$

$$F^{19}_9=\{[4\alpha]+(\overline{p+n}+n)\}.$$

$$Ne^{20}_{10}=[5\alpha]\,;\quad Ne^{21}_{10}=\{[5\alpha]+(n)\}\,;\quad Ne^{22}_{10}=\{[5\alpha]+(n+n)\}.$$

$$Na^{23}_{11}=\{[5\alpha]+(\overline{p+n}+n)\}.$$

$$Mg^{24}_{12}=[6\alpha]\,;\quad Mg^{25}_{12}=\{[6\alpha]+(n)\}\,;\quad Mg^{26}_{12}=\{[6\alpha]+(n+n)\}.$$

$$Al^{27}_{13}=\{[6\alpha]+(\overline{p+n}+n)\}.$$

$$Si^{28}_{14}=[7\alpha]\,;\quad Si^{29}_{14}=\{[7\alpha]+(n)\}\,;\quad Si^{30}_{14}=\{[7\alpha]+(n+n)\}.$$

$$P^{31}_{15}=\{[7\alpha]+(p+n+n)\}.$$

$$S^{32}_{16}=[8\alpha]\,;\quad S^{33}_{16}=\{[8\alpha]+(n)\}\,;\quad S^{34}_{16}=\{[8\alpha]+(n+n)\}.$$

Employing the hypothesis suggested, that the components outside the core are less tightly bound and that $\overline{p+n}$ is a stable unit, we may account for all neutron

emission under α-particle bombardment by two types of reaction. Thus, for beryllium, neon and magnesium * the nuclear disintegrations may be represented by

$$\mathrm{Be}_4^9 \quad \{[2\alpha]+(n)\}+\alpha \rightarrow [3\alpha]+\overset{\nearrow}{n} \quad \mathrm{C}_6^{12},$$

$$\mathrm{Ne}_{10}^{21} \quad \{[5\alpha]+(n)\}+\alpha \rightarrow [6\alpha]+\overset{\nearrow}{n} \quad \mathrm{Mg}_{12}^{24},$$

$$\mathrm{Mg}_{11}^{25} \quad \{[6\alpha]+(n)\}+\alpha \rightarrow [7\alpha]+\overset{\nearrow}{n} \quad \mathrm{Si}_{14}^{28},$$

and by similar actions neutrons might be expected as a result of the disintegration by α-particles of C_6^{13}, O_8^{17}, Si_{14}^{29}, and S_{16}^{33}. It is doubtful, however, whether these isotopes are in sufficient abundance to allow the detection of the neutrons even if they are produced.

The second type of reaction supports the hypothesis of the comparative stability of the deuton. Thus we have

$$\mathrm{Li}_3^7 \quad \{[\alpha]+(\overline{p+n}+n)\}+\alpha \rightarrow \{[2\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n} \quad \mathrm{B}_5^{10},$$

$$\mathrm{Be}_5^{11} \quad \{[2\alpha]+(\overline{p+n}+n)\}+\alpha \rightarrow \{[3\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n} \quad \mathrm{N}_7^{14},$$

$$\mathrm{F}_9^{19} \quad \{[4\alpha]+(\overline{p+n}+n)\}+\alpha \rightarrow \{[5\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n} \quad \mathrm{Na}_{11}^{22}\dagger,$$

$$\mathrm{Na}_{11}^{23} \quad \{[5\alpha]+(\overline{p+n}+n)\}+\alpha \rightarrow \{[6\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n} \quad \mathrm{Al}_{13}^{26}\dagger,$$

$$\mathrm{Al}_{13}^{27} \quad \{[6\alpha]+(\overline{p+n}+n)\}+\alpha \rightarrow \{[7\alpha]+(\overline{p+n})\}+\overset{\nearrow}{n} \quad \mathrm{P}_{15}^{30}\dagger,$$

and we might reasonably expect that by similar actions N_7^{15} and P_{15}^{31} should emit neutrons under α-particle bombardment.

In consequence of the stability of the deuton, we should expect a proton and a neutron within the nucleus to unite and form a deuton, and further, as the α-particle is a still stabler unit, we should expect that two deutons should unite to form an α-particle. Evidence in favour of this hypothesis is given by the following reactions.

First consider the disintegration of Li_3^7 by protons

$$\mathrm{Li}_3^7 \quad \{[\alpha]+(\overline{p+n}+n)\}+p \rightarrow \{[\alpha]+(\overline{p+n}+\overline{p+n})\} \rightarrow \overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}.$$

* Chadwick, Bakerian Lecture, 1933. Proc. Roy. Soc. A, cxlii. p. 1.
† Chadwick, *loc. cit.*

In this case, as the newly-formed α-particle has the same mass as the remaining core, the recoil velocity of the latter is equal to that of the ejected α-particle, so that two α-particles are emitted as if the nucleus $[2\alpha]$ were unstable. This is supported by the fact that $Be^8_4\,[2\alpha]$ is not an abundant isotope of beryllium, for in general small abundance indicates comparative instability.

This is also shown by the disintegration of boron by protons

$$B^{11}_5 \quad \{[2\alpha]+(\overline{p+n}+n)\}+p \rightarrow$$
$$\{[2\alpha]+(p+n+p+n)\} \rightarrow \overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}.$$

In this reaction we see that the energy due to the formation of the α-particle is sufficient to disrupt $[2\alpha]$ into its components. It is significant that beryllium when bombarded by protons or deutons emits α-particles of identical range. It seems probable that the unstable nucleus Be^8_4 breaks up under the bombardment and becomes two α-particles.

Similarly we have

$$F^{19}_9 \quad \{[4\alpha]+(\overline{p+n}+n)\}+p \rightarrow$$
$$\{[4\alpha]+(p+n+p+n)\} \rightarrow [4\alpha]+\overset{\nearrow}{\alpha} \quad O^{16}_8,$$

$$Na^{23}_{11} \quad \{[5\alpha]+(\overline{p+n}+n)\}+p \rightarrow$$
$$\{[5\alpha]+(p+n+p+n)\} \rightarrow [5\alpha]+\overset{\nearrow}{\alpha} \quad Ne^{20}_{10},$$

$$Al^{27}_{13} \quad \{[6\alpha]+(\overline{p+n}+n)\}+p \rightarrow$$
$$\{[6\alpha]+(p+n+p+n)\} \rightarrow [6\alpha]+\overset{\nearrow}{\alpha} \quad Mg^{24}_{12},$$

and we should therefore expect that α-particles would be obtained by the proton disintegration of nitrogen N^{15}_7 and phosphorus P^{31}_{15} by the reactions

$$N^{15}_7 \quad \{[3\alpha]+(p+n+n)\}+p \rightarrow$$
$$\{[3\alpha]+(\overline{p+n}+p+n)\} \rightarrow [3\alpha]+\overset{\nearrow}{\alpha} \quad C^{12}_6,$$

$$P^{31}_{15} \quad \{[7\alpha]+(\overline{p+n}+n)\}+p \rightarrow$$
$$\{[7\alpha]+(p+n+\overline{p+n})\} \rightarrow [7\alpha]+\overset{\nearrow}{\alpha} \quad Si^{28}_{14}.$$

Moreover, since the deuton itself is stable, it should, when its energy is sufficiently high, be capable of entering

the nucleus and interacting with the nuclear components outside the core. As a result we should expect α-particles to be emitted by nuclei of the type $\{[N\alpha]+(p+n)\}$ when bombarded with deutons, and neutrons to be emitted by nuclei of the type $\{[N\alpha]+(\overline{p+n}+n)\}$.

Evidence of such reactions are given by

$$\mathrm{Li}_3^6 \quad \{[\alpha]+(\overline{p+n})\}+\overline{p+n} \rightarrow \{[\alpha]+(\overline{p+n}+\overline{p+n})\} \rightarrow \overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha},$$

and we should expect boron B_5^{10} to disintegrate somewhat similarly, thus:

$$\mathrm{B}_5^{10} \quad \{[2\alpha]+(\overline{p+n})\}+\overline{p+n} \rightarrow \{[2\alpha]+(\overline{p+n}+\overline{p+n})\} \rightarrow \overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha},$$

since we have already shown that $[2\alpha]$ is not a very stable nucleus. Lewis, Livingston, and Lawrence * have obtained α-particles by bombarding Be_2O_3 with deutons.

In addition, α-particles were obtained from nitrogen:

$$\mathrm{N}_7^{14} \quad \{[3\alpha]+(\overline{p+n})\}+\overline{p+n} \rightarrow \{[3\alpha]+(\overline{p+n}+\overline{p+n})\} \rightarrow [3\alpha]+\overset{\nearrow}{\alpha} \quad \mathrm{C}_6^{12}.$$

Also with Li_3^7 we have

$$\mathrm{Li}_3^7 \quad \{[\alpha]+(\overline{p+n}+n)\}+\overline{p+n} \rightarrow \{[\alpha]+(\overline{p+n}+\overline{p+n})\}+n \rightarrow \overset{\nearrow}{\alpha}+\overset{\nearrow}{\alpha}+\overset{\nearrow}{n};$$

and, finally, the case of beryllium, which emits neutrons when bombarded by deutons, thus:

$$\mathrm{Be}_4^9 \quad \{[2\alpha]+(n)\}+\overline{p+n} \rightarrow \{[2\alpha]+(p+n)\}+\overset{\nearrow}{n} \quad \mathrm{B}_5^{10}.$$

In addition, as it has been shown that the deuton itself disintegrates on impact with heavier nuclei if its kinetic energy is sufficiently high, we might expect the heavy isotype of hydrogen to disintegrate into neutrons and protons if bombarded by α-particles of high energy.

Moreover, the stability of Be_4^9, O_8^{17}, and other nuclei of the form $\{[N\alpha]+(n)\}$, together with the fact that the deuton splits up into a proton and neutron, allows for the disintegration of such nuclei as B_5^{10} and N_7^{14}, with the emission of protons when bombarded by α-particles. In these cases we have to assume that the energy-change due to

* Lewis, Livingston, and Lawrence, *loc. cit.*

the entry of the α-particle into the core is sufficient to cause the deuton to break up, the proton being emitted and the neutron remaining in the nucleus. It is significant that in the scheme proposed no nuclei exist in which a proton is not accompanied by a neutron. It seems that such nuclei are unstable, so that in a disintegration of the type mentioned the neutron remains in the nucleus as forming the more stable configuration. Thus we have

$$B^{10}_{5} \quad \{[2\alpha]+(p+n)\}+\alpha \rightarrow \{[3\alpha]+(n)\}+\overset{\nearrow}{p} \quad C^{13}_{6},$$

$$N^{14}_{7} \quad \{[3\alpha]+(p+n)\}+\alpha \rightarrow \{[4\alpha]+(n)\}+\overset{\nearrow}{p} \quad O^{17}_{8}.$$

The other proton-emitting elements in the range considered are fluorine, sodium, aluminium, and phosphorus, and as far as is known these elements consist of one isotope each, these being F^{19}_{9}, Na^{23}_{11}, Al^{27}_{13}, and P^{31}_{15}. Of these, the first three are known to emit neutrons, and to explain this emission reactions of the form

$$Al^{27}_{13} \quad \{[6\alpha]+(p+n+n)\}+\alpha \rightarrow \{[7\alpha]+(p+n)\}+\overset{\nearrow}{n} \quad P^{30}_{15}, \text{ etc.},$$

have been used, the resulting nuclei in the three cases being Na^{22}_{11}, Al^{26}_{13}, and P^{30}_{15}. Is it possible to explain both the proton and neutron emission by the same isotopes?

In this connexion two points are of importance:—

1. The nuclei formed by the reactions are not found as isotopes in the analysis of sodium, aluminium, and phosphorus in the mass spectrograph, whereas stable isotopes Ne^{22}_{10}, Mg^{26}_{12}, and Si^{30}_{14} are known to exist.

2. Curie and Joliot * have obtained photographs of positrons accompanying the neutrons of aluminium, which they suggest result from the proton emission, the proton splitting up into a neutron and positron. In this case (and, indeed, with the other examples too) it seems that the instability of the resultant nuclei favours the emission of the proton, which may be interpreted by the break-up of the deuton. We conclude that for these nuclei the neutron is more tightly bound than the deuton, which splits up under α-particle bombardment with the emission of the proton. However, as neutrons are emitted, we must also conclude that the reactions mentioned do take

* Curie and Joliot, *Comptes Rendus*, cxcvi. p. 1885 (1933).

place, and that the unstable nuclei N^{22}_{10}, Mg^{26}_{12}, and P^{30}_{15} result

In addition, γ-radiation is emitted accompanying the neutrons. As the nucleus P^{30}_{15} is unstable, the γ-radiation is probably emitted by it, becoming by interaction the stable nucleus Si^{30}_{14}. Thus, if the γ-ray quantum accompanying the neutron emission interacts with the nuclear field of P^{30}_{15}, it may split up into a pair of electrons according to the Dirac Theory of the Electron. If such a pair of electrons is formed near the nucleus the positron may be emitted, the negative electron interacting with the proton in the nucleus and producing a quantum of radiation and a neutron both the electron, and the positron of the proton disappearing to form γ-radiation. As a result, the stable nucleus Si^{30}_{14} is produced and the neutron emission is accompanied by positrons.

We should, therefore, expect positrons to accompany the neutron emission of fluorine and sodium, for in these nuclei too the normal mode of disintegration under α-particle bombardment appears to be such that protons are emitted.

However, in spite of these difficulties, it seems that, on the whole, the disintegration experiments support the view that the deuton plays an important part in nuclear architecture.

This work was carried out in the Physics Department of University College, Exeter, under the direction of Professor Newman.

LXVI. *Additional Note on an Analogy between the Slow Motions of a Viscous Fluid in Two Dimensions and Systems of Plane Stress.* By J. N. Goodier, *M.A., Ph.D.**

IN a previous paper [1] it was shown that in a two-dimensional stress-distribution, represented by a bi-harmonic stress function ϕ, the boundary conditions at a stress-free hole are that $\partial\phi/\partial x$ and $\partial\phi/\partial y$ should be constants round the periphery of the hole. At another stress-free hole $\partial\phi/\partial x$, $\partial\phi/\partial y$ would have constant values different from those for the first hole, because in

* Communicated by the Author.

general there would be a resultant force transmitted between the two. The same would be true of a hole and any other stress-free boundary.

If the bi-harmonic function ϕ is now interpreted as a stream function representing a slow motion of a viscous fluid in two dimensions, the closed boundaries which are stress-free holes in the elastic plate are non-rotating rigid cylindrical boundaries in the fluid, and their different values of $\partial\phi/\partial x$, $\partial\phi/\partial y$ imply relative velocities. This will be so whether or not the elastic plate is free from dislocations, but clearly freedom from dislocations will involve some restrictions on these relative velocities. We shall prove the following theorem, restriction to steady motion being understood.

If a stressed perforated plate is free from dislocations, then in the analogous fluid motion each cylindrical boundary which corresponds to a stress-free hole in the plate moves, without rotation, in such a way that the fluid exerts no resultant force on it.

The elastic displacements in the plate have cartesian components

$$-\frac{1+\sigma}{\mathrm{E}}\cdot\frac{\partial\phi}{\partial x}+\frac{1}{\mathrm{E}}\frac{\partial f}{\partial y} \quad \text{and} \quad -\frac{1+\sigma}{\mathrm{E}}\cdot\frac{\partial\phi}{\partial d}+\frac{1}{\mathrm{E}}\frac{\partial f}{\partial x}.$$

where f is sufficiently defined by

$$\frac{\partial^2 f}{\partial x\partial y}=0, \quad \nabla^2 f=0.$$

The analytical condition for freedom from dislocations is that the displacements should be single-valued functions of the coordinates.

For any closed circuit within which no resultant force is applied, equations (3) and (4) of the earlier paper give

$$\int\frac{d}{ds}\left(\frac{\partial\phi}{\partial x}\right)ds=0, \quad \int\frac{d}{ds}\left(\frac{\partial\phi}{\partial y}\right)ds=0,$$

and these imply that $\partial\phi/\partial x$, $\partial\phi/\partial y$ are single-valued. The elastic displacements will therefore be single-valued if $\partial f/\partial x$ and $\partial f/\partial y$ are single-valued—that is, if

$$\int\frac{d}{ds}\left(\frac{\partial f}{\partial x}\right)dx=0, \quad \int\frac{d}{dx}\left(\frac{\partial f}{\partial y}\right)ds=0,$$

the integrals being taken round any circuit enclosing an unloaded hole. Expanding the integrands, these conditions become

$$\left.\begin{aligned}\int\left(\frac{\partial^2 f}{\partial x^2}\cdot\frac{dx}{ds}+\frac{\partial^2 f}{\partial x\partial y}\frac{dy}{ds}\right)ds=0,\\ \int\left(\frac{\partial^2 f}{\partial x\partial y}\cdot\frac{dx}{ds}+\frac{\partial^2 f}{\partial y^2}\cdot\frac{dy}{ds}\right)ds=0.\end{aligned}\right\}\quad . \quad . \quad . \quad (1)$$

By definition $\partial^2 f/\partial x\,\partial y=\nabla^2\phi$, so that

$$\frac{\partial}{\partial x}\nabla^2\phi=\frac{\partial}{\partial y}\frac{\partial^2 f}{\partial x^2},\quad \frac{\partial}{\partial y}\nabla^2\phi=\frac{\partial}{\partial x}\cdot\frac{\partial^2 f}{\partial y^2}=-\frac{\partial}{\partial x}\frac{\partial^2 f}{\partial x^2},$$

f being a harmonic function. Thus $\partial^2 f/\partial x^2$, which will now be denoted by Ω, is the function conjugate to $\nabla^2\phi$. The conditions (1) can now be written

$$\int\left(\Omega\frac{dx}{ds}+\nabla^2\phi\frac{dy}{ds}\right)ds=0,\quad \int\left(\nabla^2\phi\frac{dx}{ds}-\Omega\frac{dy}{ds}\right)ds=0 \quad (2)$$

In particular, they will hold for a circuit coinciding with the periphery of a stress-free hole.

Consider, now, the analogous fluid motion. The forces on a cylindrical boundary are due to the normal stress, which is the same in magnitude as the mean pressure p at any rigid boundary, and to the shearing stress μq, where q is the rate of shearing strain. It may be shown by means of the general formula for $q^{(1)}$, and the conditions

$$\frac{d}{ds}\left(\frac{\partial\phi}{\partial x}\right)=0,\quad \frac{d}{ds}\left(\frac{\partial\phi}{\partial y}\right)=0,$$

that at a non-rotating rigid boundary $q=-\nabla^2\phi$. We also have $p=\mu\Omega$. The cartesian components of the force exerted by the fluid on a cylindrical boundary are

$$\int\left(\mu q\frac{dx}{ds}+p\frac{dy}{ds}\right)ds \quad\text{and}\quad \int\left(-p\frac{dx}{ds}+\mu q\frac{dy}{ds}\right)ds,$$

or

$$\mu\int\left(-\nabla^2\phi\frac{dx}{ds}+\Omega\frac{dy}{ds}\right)ds \quad\text{and}\quad \mu\int\left(-\Omega\frac{dx}{ds}-\nabla^2\phi\frac{dy}{ds}\right)ds,$$

the integrals being taken round the boundary. The condition (2) expresses the vanishing of these components, leading to the result already stated.

In the earlier paper[1] it was stated that pure bending of a strip with a hole corresponds to translation of a cylinder in a stream. The theorem just proved indicates that the velocity of the cylinder is that acquired when it is subject to no constraints except the couple required to prevent rotation. Thus Howland's solution (see ref. 1) for bending of a strip with a central hole provides the stream function for a stream between parallel walls in which a circular cylinder is drifting freely midway between them, no couple being required on account of symmetry.

Solutions of the elastic problem have also been obtained for deep beams with elliptical holes on the neutral axis[2]. When the hole has one axis along the neutral axis, we have the analogue of a stream in which an elliptical cylinder is drifting.

Some types of motion for a cylinder near a plane wall, when the prevailing flow is laminar, are provided by Jeffery's solution of the stress-problem for a perforated semi-infinite plate subject to tension parallel to the edge[3]. On account of symmetry the shear stress vanishes on the normal to the edge through the centre of the hole, and this implies that in the analogous fluid motion there is no velocity of approach. There is, however, a motion of the cylinder parallel to the plane, the latter being supposed fixed. The above theorem shows that it is the motion ensuing when the cylinder is free from resultant force but is prevented from rotating by a couple.

It is known that a couple applied to a cylinder near a plane wall will produce rotation without any tendency to translation[4]. Thus, if a couple is applied to the cylinder in the laminar flow, equal and opposite to that which is preventing rotation, the cylinder becomes free. It acquires an angular velocity, but its motion of translation is unchanged, and we have the result that a (circular) cylinder adrift near a plane wall can move parallel to the wall.

References.

(1) J. N. Goodier, Phil. Mag. xvii. p. 554 (1934).
(2) K. Wolf, *Zeitschrift für Technische Physik*, no. 8 (1921).
(3) G. B. Jeffery, Phil. Trans. A, 221, p. 288 (1920).
(4) R. A. Frazer, Phil. Trans. A, 225, p. 93 (1925).

Ontario Research Foundation.
December 12th, 1933.

LXVII. *On the Application of Least Squares.*—II. *By* W. Edwards Deming, *Ph.D., Associate Physicist, Bureau of Chemistry and Soils, U.S. Department of Agriculture, Washington, D.C.**

Abstract.

A solution of the general problem in least squares, together with examples of its applications to several special cases, was given in a previous paper. The author now presents further illustrations of the adjustment of observations and determination of parameters, together with suggestions for systematic procedure in computation. The problems here treated are the logarithmic decrement of a balance, the exponential law, and the laws represented by the equations $ya^x=b$, $yz^x=b$, and $yz^x=w$, wherein a and b are parameters, and x, y, z, w are observed coordinates. The last three equations are respectively the ones needed in the determination of e and h (electronic charge and Planck's constant) under the three situations: (a) neither e nor h directly observed, (b) direct observations on e included, (c) direct observations on both e and h included. Provisions in the general solution allow any relation between e and h to be forced. In particular, the normal equations are set up and illustrated with Eddington's relation $hc/2\pi e^2=137$ forced or released at will, using Birge's weighting. The disparity between the two pairs of values of e and h that are obtained with and without forcing any theoretical relation constitutes a basis for accepting or rejecting the theory.

In the first paper the question was raised concerning how the absolute values of the weights of the y and z coordinates of n points could be ascertained when the weight of any one of the x coordinates has been arbitrarily assigned. Only one of the $3n$ weights is arbitrary; the others are all relative to it. It is now clear that on account of the manner in which the weights enter the solution, any weight w_i must be taken as C/l_i^2, where l_i is the probable error of the ith observation, and C is an arbitrary constant that retains the same value for all coordinates, regardless of their dimensions and units. The solution is independent of the choice of C and of the scale chosen

* Communicated by H. F. Stimson, Ph.D.

along any coordinate axis; each weight w_i is used with a derivative that automatically cancels the effect of a change in units.

The procedure for computation in curve fitting is uniform and methodical, irrespective of the empirical formula, and the results are independent of the form in which it is written.

Introduction.

IN an earlier paper * I derived equations for the adjustment by least squares of observations connected by conditions that may or may not involve unknown parameters. Provision was made for the fact that the parameters may not all be independent. From this development it is clear that the solution of any problem by least squares yields not only the "most probable" values of the parameters that may enter, but also yields the adjusted values of all the observed quantities, and the solution is unique.

Karl Pearson's paper † "On Lines and Planes of closest Fit to Systems of Points in Space," published in 1902, first called attention to the fact that the usual custom of throwing the adjustment all on to one coordinate may be inadvisable in many physical problems. Pearson solved neatly the case where the relations between the coordinates are linear and where the observations on all coordinates are of equal weight. His work seems to have remained unnoticed by physicists, though it is common property now among statisticians. Probably one reason for this is that Pearson's solution requires calculations of the correlation coefficients—a term indispensable to statisticians, but little used by physicists. In 1920 R. Meldrum Stewart ‡ indicated a possible method for obtaining the solution of the general problem, and he illustrated it with two special examples. In 1921 Lowell J. Reed § discussed Pearson's results for two dimensions and described an ingenious instrument for

* W. Edwards Deming, Phil. Mag. xi. pp. 146–158 (January 1931).
† Karl Pearson, Phil. Mag. ii. pp. 559–572 (1902).
‡ R. Meldrum Stewart, Phil. Mag. xl. pp. 217–227 (1920).
§ Lowell J. Reed, Metron, i. no. 3, pp. 54–61 (1921).

fitting a line to points in a plane, giving the coordinates equal weight. In the same year Corrado Gini * derived equations for the slope and intercept of a line when the x coordinates all have equal weight and the y coordinates all have equal weight. H. S. Uhler's discussion † of the same problem in 1923 went far in clearing up misunderstandings of the principle of least squares. In 1930 the writer ‡ used Lagrange multipliers for the linear empirical equation, allowing all coordinates to have arbitrary weights, and in 1931 § he extended the same scheme to the general problem. All the previous results, at least to the first order of the corrections, are contained in this general solution as special cases.

Frequent misuses of least squares and the concomitant confusion, so well analysed by Stewart and Uhler (*loc. cit.*), are an inevitable result of the faulty approach to the subject made by most (if not all) writers of treatises on the adjustment of observations. For this reason it seems desirable that the principle of least squares should be carefully explained and more elaborately illustrated. To this end, the present communication will expand that portion of the earlier paper which dealt with applications of the general solution to special problems, and so will provide further examples of rigorous treatment.

A secondary purpose of writing this article is to answer a question propounded in the first paper regarding the assignment of weights to the several coordinates of the points, when the coordinates are different dimensionally. The discussion on this matter will be deferred to the last section.

Notation and the General Solution.

Since little more effort is required to follow the solution of the general problem than is required to deduce the normal equations for most special cases *ab initio*, it seems advisable simply to write each example in the general notation and substitute directly into the equations of the previous article It will become obvious that the correct

* Corrado Gini, Metron, i. no. 3, pp. 63–82 (1921).

† H. S. Uhler, J. Opt. Soc. and Rev. Sci. Inst. vii. pp. 1043–1066 (1923).

‡ W. Edwards Deming, Proc. London Phys. Soc. xlii. pp. 97–107 (1930).

§ *Loc. cit.*

application of least squares to any problem is straightforward and methodical—usually more so than a faulty adjustment.

The notation of the previous paper will be used, and the equations when duplicated will be given their old numbers. The notation has been developed to take care of the most general circumstances. The observations may or may not be on the coordinates of points; in any case x_i will denote the ith observation. If the observations are on coordinates, i will run from 1 to $2n$ or from 1 to $3n$, depending on whether each of the n points has two or three coordinates. The letter q will be used for the number of dimensions. In practice q is usually 2, but it is occasionally 3, and may be higher. The observed x, y, z coordinates will sometimes be denoted by x_h, y_h, z_h, but more often, for compactness in summation and writing, by x_h, x_{n+h}, x_{2n+h}. r_i will be used to designate the adjusted value of x_i: accordingly r_h, r_{n+h}, r_{2n+h} will be the adjusted values of the coordinates that were observed to be x_h, x_{n+h}, x_{2n+h}. w_i will be the weight of x_i.

In the relations connecting the adjusted values there may occur certain unknown constants or parameters, values for which are to be found—these will be designated by $a, b, c, \ldots, p$. Approximate values, however arrived at, will be denoted by $a_0, b_0, c_0, \ldots, p_0$. These relations, m in number, existing among the adjusted values may be expressed by the m equations

$$F^h(r_1, r_2, \ldots, r_{qn}; \quad a, b, \ldots, p)=0, \quad h=1, 2, \ldots, m. \quad (1)$$

The parameters may not all be independent, *e. g.*, there may exist certain relations between two or more of them, or the empirical curve (or surface) may be known to pass rigorously through a particular point or points; these facts can be expressed by l further relations:

$$F^h(a, b, \ldots, p)=0, \quad h=m+1, m+2, \ldots, m+l. \quad (2)$$

It is convenient to let $F^1_0, F^2_0, \ldots$ designate the left-hand sides of eqs. (1) and (2) when the observations $x_1, x_2, \ldots, x_{qn}$ and the approximate parameters $a_0, b_0, \ldots, p_0$ are substituted therein. $F^1_0, F^2_0, \ldots$ will be small quantities that signify how much the observed and approximate values fail to satisfy the requirements that

are to be imposed on the adjusted values. The derivatives are denoted by subscripts :

$$F^h_i \equiv \frac{d}{dx_i} F^h(x_1, x_2, \ldots, x_{qn};\ a_0, b_0, \ldots, p_0),$$

$$i=1, 2, \ldots, qn;\ h=1, 2, \ldots, m+l,$$

$$F^h_a \equiv \frac{d}{da_0} F^h(x_1, x_2, \ldots, x_{qn};\ a_0, b_0, \ldots, p_0),$$

$$F^h_b \equiv \frac{d}{db_0} F^h(x_1, x_2, \ldots, x_{qn};\ a_0, b_0, \ldots, p_0), \text{ etc.}$$

It is the corrections to the observations and to the approximate parameters that are to be determined. These can be denoted by $r_i - x_i$, $a-a_0$, $b-b_0, \ldots$, or, more conveniently, by δx_i, δa, $\delta b, \ldots$ If the corrections are not too large (and they will not be if the data are worth adjusting, and if some care has been taken in selecting the approximate values of the parameters), the conditions can be replaced closely enough with equations linear in the corrections by using Taylor's series and dropping powers higher than the first. Then in place of eqs. (1) and (2) we may write

$$\Sigma_i F^h_i \delta x_i + F^h_a \delta a + F^h_b \delta b + \ldots = -F^h_0, \quad h=1, 2, \ldots, m. \quad (3)$$

$$F^h_a \delta a + F^h_b \delta b + \ldots = -F^h_0, \quad h=m+1, m+2, \ldots, m+l. \quad (4)$$

The principle of least squares requires the sum of the weighted squares of the residuals to be as small as possible consistent with the conditions of the problem, *i. e.*,

$$\phi \equiv \sum_1^{qn} w_i (r_i - x_i)^2 = \sum_1^{qn} w_i \delta x_i^2$$

is to be a minimum with respect to the adjusted values $r_1, r_2, \ldots, r_{qn}$, a, b, $c, \ldots, p$, which are subject to the conditions (1) and (2), or (3) and (4). The problem, then, is one in the maxima and minima of implicit functions, for which the method of Lagrange * is convenient. The solution, which was obtained in the first paper, is contained in the following symmetrical equations—the "general normal equations." $\lambda_1, \lambda_2, \ldots$ are Lagrange multipliers.

* Lagrange, 'Méchanique Analytique,' i. p. 74. Given also in Benjamin Williamson's 'Differential Calculus,' ch. xi. (Longmans, 1899), and in most treatises on mechanics.

$$\left.\begin{array}{cccccccccccc}
\lambda_1 & \lambda_2 & \ldots & \lambda_m & \lambda_{m+1} & \lambda_{m+2} & \delta a & \delta b & \delta c & \ldots & \text{Const.} \\
\hline
L_{11} & L_{12} & \ldots & L_{1m} & 0 & 0 & -F^1_a & -F^1_b & -F^1_c & \ldots & =F^1_0 \\
 & L_{22} & \ldots & L_{2m} & 0 & 0 & -F^2_a & -F^2_b & -F^2_c & \ldots & =F^2_0 \\
 & & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
 & & & L_{mm} & 0 & 0 & -F^m_a & -F^m_b & -F^m_c & \ldots & =F^m_0 \\
 & & & & 0 & 0 & -F^{m+1}_a & -F^{m+1}_b & -F^{m+1}_c & \ldots & =F^{m+1}_0 \\
 & & & & & 0 & -F^{m+2}_a & -F^{m+2}_b & -F^{m+2}_c & \ldots & =F^{m+2}_0 \\
 & & & & & & 0 & 0 & 0 & \ldots & =0 \\
 & & & & & & & 0 & 0 & \ldots & =0 \\
 & & & & & & & & 0 & \ldots & =0 \\
 & & & & & & & & \vdots & & \vdots
\end{array}\right\}, \quad (7)$$

For brevity, only the coefficients of the unknowns $\lambda_1, \lambda_2, \ldots, \lambda_{m+2}, \delta a, \delta b, \delta c, \ldots$ are tabled; thus the first row represents the equation *

$$L_{11}\lambda_1+L_{12}\lambda_2+\ldots+L_{1m}\lambda_m+0\cdot\lambda_{m+1}+0\cdot\lambda_{m+2}$$
$$-F^1_a\delta a-F^1_b\delta b-F^1_c\delta c-\ldots=F^1_0.$$

Herein

$$L_{rs}\equiv\sum_1^{qn}F^r_iF^s_i/w_i=L_{sr}.$$

The coefficients L_{rs} are called "weighting factors."

On account of the symmetry about the diagonal in this table, only the upper part is written. The portion below the diagonal is not used in the abridged method for the solution of symmetrical equations †. The coefficients for the third equation, for example, are found by reading down the third column as far as the diagonal, then proceeding across.

The corrections $\delta a, \delta b, \ldots$ are found directly from the normal equations. The corrections to the observations,

$$\delta x_i=\frac{1}{w_i}\sum_1^m \lambda_k F^k_i, \quad i=1, 2, \ldots, qn, \quad . \quad . \quad . \quad (6)$$

* In the 1931 paper, the right-hand side of this equation was inadvertently written $-F^1_0$ instead of $+F^1_0$.

† See, for example, O. M. Leland, 'Practical Least Squares,' Art. 60 (McGraw-Hill, 1921).

can also be computed, if desired, after the solution of the normal equations has yielded $\lambda_1, \lambda_2, \ldots, \lambda_m$.

It is important to realize that the method of least squares is a process that takes account of the weight of each observation; in fact, it is essentially a method for the *adjustment of observations.* Values for the parameters (if any) are obtained as a by-product. It often happens that this by-product is the most important part of the problem. When this is the case and only the parameters are desired, the solution of the normal equations can be stopped as soon as δa, δb, $\delta c, \ldots$ have been found; it is not necessary to continue the solution to determine $\lambda_1, \lambda_2, \lambda_3, \ldots$ and find the corrections to the observations. Nevertheless, the principle behind the process is always to minimize the weighted squares of the corrections to the observations, and the complete result can in every case be obtained by continuing the solution.

There are six or more schemes in common use for *finding the parameters* in an empirical formula, and they will all give different results with a set of actual data. It cannot be said that any one of them always gives the best parameters. Further, the relative amounts of computation that are required vary with the problem, and the amount of labour that can be justified depends on the circumstances; hence no general rule can be laid down as to which process should be used for finding parameters. In spite of the importance of inquiries into the relative merits that each method possesses under particular circumstances, I must again omit all such discussion in order that these writings may be purely expositions of the principle of least squares.

Illustrations with various types of condition equations follow.

Conditions without Parameters.

When no parameters are contained in the conditions, the general normal equations reduce to

$$\left.\begin{aligned} L_{11}\lambda_1+L_{12}\lambda_2+\ldots+L_{1m}\lambda_m&=F_0^1,\\ L_{22}\lambda_2+\ldots+L_{2m}\lambda_m&=F_0^2,\\ \vdots\qquad&\quad\vdots\\ L_{mm}\lambda_m&=F_0^m.\end{aligned}\right\}\quad . \quad . \quad . \quad (8)$$

The corrections δx_i are, of course, found from eqs. (6) after eqs. (8) have been solved for the multipliers $\lambda_1, \lambda_2, \ldots$

λ_m. This type of problem is adequately treated in many texts and requires no amplification here. A simple example is the three angles of a plane triangle; the sum of their adjusted values must be 180°. Another example is the determination of e and h when both e and h are directly and indirectly observed (*vide infra*).

Conditions that contain Parameters and involve Quantities not Coordinates.

It is not generally realized that an empirical formula imposes conditions on the adjusted values of observed quantities in the same way that geometrical requirements impose conditions. The principle of least squares is often remoulded to fit each new problem, and the result is confusion. The *principle* remains the same; the only difference between one type of problem and another lies in the condition equations.

An example of observations on quantities that are not considered to be coordinates, but which are related through parameters, is the swing of the pointer on a balance where successive arcs are supposed to be related by the "logarithmic decrement"*. Let y_h denote the observed amplitude for the hth swing; then if the readings were devoid of error they would satisfy the empirical formula

$$ln|y_h-y_{h-1}|=b-ah,\ h=2,\ 3,\ 4,\ \ldots,\ n, \quad . \quad (16)$$

wherein a is the logarithmic decrement. The adjusted values of $y_1,\ y_2,\ \ldots,\ y_n,\ a,\ b$, must satisfy these $n-1$ equations exactly; hence the empirical formula imposes $n-1$ conditions on these $n+2$ quantities.

The general solution can be made immediately applicable by writing

$$\mathrm{F}^h(r_1,\ r_2,\ \ldots,\ r_n;\ a,\ b)\equiv ln\,|r_h-r_{h-1}|-b+ah,$$
$$h=2,\ 3,\ \ldots,\ n. \quad . \quad . \quad (17)$$

Here $r_1,\ r_2,\ \ldots$ are the adjusted values of the amplitudes. The adjusted values of $r_1,\ r_2,\ \ldots,\ r_{2n},\ a,\ b$ will make these $n-1$ expressions vanish.

We shall assume that the swings are observed with equal precision; then the w_h are all equal and may be used as unity for convenience.

$$\mathrm{F}_0^h=ln\,|y_h-y_{h-1}|+a_0h-b_0,\ h=2,\ 3,\ \ldots,\ n. \quad (18)$$

* See, for example, W. Felgentraeger, 'Feine Waagen Wägungen und Gewichte,' pp. 24–32, pp. 242–245 (Julius Springer, 1932).

F^1_0, F^2_0, . . . differ from zero only to the extent that the observed swings and the approximate values a_0 and b_0 fail to make the expressions (17) vanish. Satisfactory values for a_0 and b_0 can be found quickly by graphing a few values of $\log|y_h-y_{h-1}|$ against h and finding the slope and intercept of a line drawn close to the points.

The various derivatives of F^h are easily found to be

$$\left.\begin{array}{l} F^h_h=\pm Y_h,\ F^h_{h-1}=\mp Y_h,\quad y_h \gtrless y_{h-1}, \\ F^h_a=h,\ F^h_b=-1, \end{array}\right\} \quad . \quad . \quad (19)$$

where

$$Y_h=1/|y_h-y_{h-1}|. \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

The weighting factors in the normal eqs. (7) will be

$$L_{rs}=\sum_{h=2}^{n} F^r_h F^s_h \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

since the weights are all unity. We have then

$$\left.\begin{array}{r} L_{hh}=F^h_{h-1}F^h_{h-1}+F^h_hF^h_h=2Y^2_h, \\ L_{h,\,h-1}=F^h_{h-1}F^{h-1}_{h-1}=Y_hY_{h-1}=L_{h-1,\,h}, \\ L_{rs}=0 \quad \text{if } |r-s|>1. \end{array}\right\} \quad . \quad . \quad . \quad (22)$$

If there were no errors of observation, Y^2_h would be equal to $Y_{h-1}Y_{h+1}$, so the equalities

$$L_{h,\,h-1}L_{h,\,h+1}=Y_{h-1}Y^2_hY_{h+1}=\tfrac{1}{4}L^2_{hh} \quad . \quad . \quad . \quad (23)$$

would be satisfied. In practice these will hold so closely that they will constitute a check on the computation. Or, if desired, by means of these relations the coefficients off the diagonal can be computed from those on the diagonal.

The normal equations for five readings are the following. Extension to more readings is easily made.

$$\left.\begin{array}{ccccccc} \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 & \delta a & \delta b & \text{Const.} \\ \hline 2Y^2_2 & Y_2Y_3 & 0 & 0 & -2 & 1 & F^1_0 \\ & 2Y^2_3 & Y_3Y_4 & 0 & -3 & 1 & F^2_0 \\ & & 2Y^2_4 & Y_4Y_5 & -4 & 1 & F^3_0 \\ & & & 2Y^2_5 & -5 & 1 & F^4_0 \\ & & & & 0 & 0 & 0 \\ & & & & & 0 & 0 \end{array}\right\} \quad . \quad (24)$$

Normal equations will always be symmetrical, so the usual short cuts and checks in formation and solution should be used*. If only the logarithmic decrement is desired, the solution can be stopped when δa has been computed. The adjusted values of a and b are $a_0+\delta a$ and $b_0+\delta b$.

If it is desired to adjust the observations it is necessary to continue the solution and find λ_2, λ_3, λ_4, λ_5. Then by eqs. (6):

$$\left.\begin{aligned}\delta y_1&=0 \qquad -\lambda_2 F_1^2= && -\lambda_2 Y_2,\\ \delta y_2&=-\lambda_2 F_2^2-\lambda_3 F_2^3= && \lambda_2 Y_2+\lambda_3 Y_3,\\ \delta y_3&=-\lambda_3 F_3^3-\lambda_4 F_3^4= && -\lambda_3 Y_3-\lambda_4 Y_4,\\ \delta y_4&=-\lambda_4 F_4^4-\lambda_5 F_4^5= && \lambda_4 Y_4+\lambda_5 Y_5,\\ \delta y_5&=-\lambda_5 F_5^5-0 \qquad = && -\lambda_5 Y_5.\end{aligned}\right\} \quad . \quad . \quad . \quad (25)$$

Check: $\Sigma \delta y_h=0$.

The signs before the λY in eqs. (25) are written on the supposition that the swings in the positive direction of the scale are given the odd numbers; otherwise these signs are to be reversed.

The adjusted values

$$y_1+\delta y_1,\ y_2+\delta y_2,\ \ldots,\ a_0+\delta a,\ b_0+\delta b$$

will satisfy the empirical formula (16) exactly.

No other possible values of δy_1, δy_2, ... can have $\sum_1^n \delta y_h^2$ as small as those calculated by eqs. (25). Hence we have found the true least squares solution.

It is interesting to solve this problem with a different approach. Write

$$F^h(r_1, r_2, \ldots, r_n;\ a, b)=ln\frac{|r_h-r_{h-1}|}{|r_{h+1}-r_h|}-a,$$

$$h=2, 3, \ldots, n-1. \quad . \quad (26)$$

Here there is only one parameter, the logarithmic decrement.

$$F_0^h=ln\frac{|y_h-y_{h-1}|}{|y_{h+1}-y_h|}-a_0. \quad . \quad . \quad . \quad . \quad . \quad (27)$$

* See, for example, O. M. Leland, 'Practical Least Squares,' Art. 60. (McGraw-Hill, 1921.)

$$\left.\begin{aligned} &F^h_a = -1 \\ &F^h_{h-1} = \mp Y_h, \ y_h \gtrless y_{h-1}, \\ &F^h_h = \pm Y_h \mp Y_{h+1}, \ y_h \gtrless y_{h-1}, \\ &F^h_{h+1} = \pm Y_{h+1}, \ y_h \gtrless y_{h-1}. \end{aligned}\right\} \quad . \ . \ . \ . \ . \quad (28)$$

$$\left.\begin{aligned} L_{hh} &= 2\{Y_h^2 - Y_h Y_{h+1} + Y_{h+1}^2\} \\ &= 2Y_h Y_{h+1} \text{ (closely enough)} \\ L_{h,h-1} &= L_{h-1,h} = Y_h\{Y_{h-1} - 2Y_h + Y_{h+1}\} = 0 \\ &\qquad\qquad \text{(closely enough),} \\ L_{h,h-2} &= L_{h-2,h} = -Y_{h-1}Y_h = \tfrac{1}{2}L_{h-1,h-1}. \end{aligned}\right\} . \quad (29)$$

All the other L_{rs} are zero.

The normal equations for five readings would appear as

$$\begin{array}{ccccc} \lambda_2 & \lambda_3 & \lambda_4 & \delta a & \text{Const.} \\ \hline 2Y_2Y_3 & 0 & -Y_3Y_4 & 1 & F_0^2 \\ & 2Y_3Y_4 & 0 & 1 & F_0^3 \\ & & 2Y_4Y_5 & 1 & F_0^4 \\ & & & 0 & 0 \end{array} \quad \Bigg\} \ . \ . \quad (30)$$

δa is found from these. The adjusted value $a_0 + \delta a$ will, of course, be the same as found by eqs.(24).

The corrections to the observations will be

$$\left.\begin{aligned} \delta y_1 &= -\lambda_2 F_1^2 + 0 \quad + 0 \quad = -\lambda_2 Y_2, \\ \delta y_2 &= -\lambda_3 F_2^3 - \lambda_2 F_2^2 + 0 \quad = \lambda_3 Y_3 - \lambda_2(Y_3 - Y_2), \\ \lambda y_3 &= -\lambda_4 F_3^4 - \lambda_3 F_3^3 - \lambda_2 F_3^2 = -\lambda_4 Y_4 + \lambda_3(Y_4 - Y_3) + \lambda_2 Y_3, \\ \delta y_4 &= \ 0 \quad -\lambda_4 F_4^4 - \lambda_3 F_4^3 = \qquad -\lambda_4(Y_5 - Y_4) - \lambda_3 Y_4, \\ \delta y_5 &= \ 0 \quad -0 \quad -\lambda_4 F_5^4 = \qquad\qquad +\lambda_4 Y_5. \end{aligned}\right\}$$

$$. \ . \ . \quad (31)$$

and these, too, will have the same numerical values as those previously found. The signs before the λY in the last equations are again written on the supposition that the swings in the positive direction are numbered 1, 3, 5, . . . ; otherwise they should be reversed.

Another example where the conditions contain parameters and involve observations that are not coordinates will be found in the section on the determination of e and h when direct observations on e are included.

Conditions that contain Parameters and involve Coordinates of Points. Curve Fitting.

Here the observations are on the coordinates of points and the points are related only through the parameters. If the observations were devoid of error, then the empirical curve, if correct and if used with the true values of the parameters, would pass through each point. Actually, of course, the observations are not devoid of error, so the observed points miss the empirical curve by varying amounts. To each *observed* point there corresponds an *adjusted* point on the empirical curve. The line segment joining the observed and adjusted positions of a point will be parallel to the y axis if the x coordinate is infallible compared with the y coordinate, and it will be parallel to the x axis if the converse is true. The segment is perpendicular to the empirical curve if the x and y coordinates have equal weight. The extension to three or more dimensions is obvious *.

For the empirical formula $F(x, y, z;\ a, b, c, \ldots)=0$ we write

$$F^h(r_1, r_2, \ldots, r_{3n};\ a, b, c, \ldots)\equiv F(r_h, r_{n+h}, r_{2n+h};\ a, b, c, \ldots),\ h=1, 2, \ldots, n;$$

then the condition equations (1) or (3) become the requirement that the adjusted values $r_1, r_2, \ldots, r_{3n}$, $a, b, c, \ldots$ satisfy the formula exactly. If there are n points there are n such conditions. In addition, there may be requirements that the curve (or surface) pass rigorously through one or more given points, which is saying that the parameters may not all be independent; these would constitute further relations between the parameters and are taken care of by eqs. (2) or (4).

Since in the problems now under consideration the conditions involve each point separately, the corrections to the observations reduce to one term each:

$$\left.\begin{aligned}\delta x_h&=-\frac{1}{w_h}\lambda_h F^h_h,\\ \delta y_h=\delta x_{n+h}&=-\frac{1}{w_{n+h}}\lambda_h F^h_{n+h},\\ \delta z_h=\delta x_{2n+h}&=-\frac{1}{w_{2n+h}}\lambda_h F^h_{2n+h}.\end{aligned}\right\} \quad \ldots \quad (10)$$

* W. Edwards Deming, Phil. Mag. xi. pp. 146–158 (January 1931).

Further, the weighting factors L_{rs} off the diagonal do not exist, and only the diagonal terms are left. They are

$$L_h = \frac{F^h_h F^h_h}{w_h} + \frac{F^h_{n+h} F^h_{n+h}}{w_{n+h}} + \frac{F^h_{2n+h} F^h_{2n+h}}{w_{2n+h}} \quad . \quad . \quad . \quad (11)$$

If any coordinate x has infinite weight, then $1/w_i = 0$ and the corresponding term disappears from L_h.

The general normal equations now can be resolved into the following two sets of equations.

$$\left.\begin{array}{cccccccc}
\lambda_1 & \lambda_2 & \ldots & \lambda_n & \delta a & \delta b & \delta c & \ldots \text{ Const.} \\
\hline
L_1 & 0 & \ldots & 0 & -F^1_a & -F^1_b & -F^1_c & \ldots = F^1_0 \\
0 & L_2 & \ldots & 0 & -F^2_a & -F^2_b & -F^2_c & \ldots = F^2_0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \ldots & L_n & -F^n_a & -F^n_b & -F^n_c & \ldots = F^n_0
\end{array}\right\} . \quad (12)$$

$$\left.\begin{array}{ccccccc}
\delta a & \delta b & \delta c & \ldots & \lambda_{n+1} & \lambda_{n+2} & \text{Const.} \\
\hline
\frac{F^h_a F^h_a}{L_h} & \frac{F^h_a F^h_b}{L_h} & \frac{F^h_a F^h_c}{L_h} & \ldots & F^{n+1}_a & F^{n+2}_a & = -\frac{F^h_a F^h_0}{L_h} \\
 & \frac{F^h_b F^h_b}{L_h} & \frac{F^h_b F^h_c}{L_h} & \ldots & F^{n+1}_b & F^{n+2}_b & = -\frac{F^h_b F^h_0}{L_h} \\
 & & \frac{F^h_c F^h_c}{L_h} & \ldots & F^{n+1}_c & F^{n+2}_c & = -\frac{F^h_c F^h_0}{L_h} \\
 & & & & \vdots & \vdots & \vdots \\
 & & & & 0 & 0 & = -F^{n+1}_0 \\
 & & & & & 0 & = -F^{n+2}_0
\end{array}\right\} \quad . \quad . \quad . \quad (13)$$

Summation over h from 1 to n is implied here in all terms where h occurs as an index.

If there are no rigorous conditions on the parameters, eqs. (2) and (4) and $\lambda_{n+1}, \lambda_{n+2}, \ldots$ do not exist, and the columns and rows that are written with F^{n+1} and F^{n+2} are to be deleted. The corrections $\delta a, \delta b, \ldots$ are determined from the symmetrical eqs. (13); the multipliers $\lambda_1, \lambda_2, \ldots$ for use in eqs. (10) are determined from eqs. (12). If the corrections to the observations are not desired, it is not necessary to set up eqs. (12) at all.

The multipliers λ_{n+1}, λ_{n+2}, . . . are of no use in themselves, but by their presence they influence the values of δa, δb, . . . to be obtained from eqs. (13), thus imposing the requirements expressed in eqs. (2) or (4).

Practical Procedure.

Probably the most practical procedure for setting up the eqs. (13) is first to tabulate in columns under the following headings :

$$h \qquad F^h_a \qquad F^h_b \qquad F^h_c \ . \ . \ . \ F^h_0 \qquad L_h \qquad \sqrt{L_h}$$

h takes on the values 1, 2, 3, . . . , n, and also $n+1$, $n+2$, . . . , $n+l$ in case there are any relations expressed by eq. (2) ; in practice, there are usually not more than one of these extra relations. There will be $n+l$ rows under each heading, except under L_h and $\sqrt{L_h}$, where there will be only n rows. Then from this table another is formed under the headings

$$h \qquad F^h_a/\sqrt{L_h} \qquad F^h_b/\sqrt{L_h} \qquad F^h_c/\sqrt{L_h} \ . \ . \ . \ F^h_0/\sqrt{L_h}.$$

There will be n rows under each of these.

The sum of the squares and cross products required for the normal eqs. (13) are to be formed from the latter table. The procedure is thus definite and uniform. It is always the same, regardless of the empirical formula, and the form in which it appears, and special requirements imposed on it. The usual checks and short cuts in the formation and solution of symmetrical equations should, of course, be used (see footnote, p. 813).

Examples.

(a) *The linear equation.*—The normal equations for determining the two parameters in the empirical formula $y=ax+b$, from observations of arbitrary weight on the x and y coordinates of n points, follow easily from eqs. (13). Provision can be made for forcing the line to pass through a given point x_0, y_0. This problem was covered in the 1931 paper and requires no further treatment. There are, however, some interesting features in regard to the linear equation that I hope to present in a later article

(b) *The exponential law.*—Here the x and y coordinates of each point, if devoid of error, would satisfy

$$y=be^{ax}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (32)$$

Approximate values of a and b can be obtained quickly by plotting log y against x. We then write

$$\mathrm{F}^h(r_1, r_2, \ldots, r_{2n};\ a, b)\equiv be^{ar_h}-r_{n+h},\ h=1, 2, \ldots, n. \ (33)$$

By giving h the value $n+1$ we can force the curve to pass through the point x_0, y_0.

$$\left.\begin{aligned} &\mathrm{F}^h_0=b_0e^{a_0x_h}-y_h, \\ &\mathrm{F}^h_a=bx_he^{ax_h}=x_hy_h.\ \mathrm{F}^h_b=e^{ax_h}\doteq y_h/b, \\ &\mathrm{F}^h_h\doteq ay_h,\ \mathrm{F}^h_{n+h}=-1, \\ &\mathrm{L}_h=a^2y_h^2/w_h+1/w_{n+h}. \end{aligned}\right\} \quad . \quad . \quad . \qquad (34)$$

In the actual computation of the derivatives and weighting factors, the approximate values of a, b, c. . . . would ordinarily be inserted. This is not strictly necessary, however, for any values close to them may be used safely. To indicate this fact, and for the sake of brevity, the subscript o will be omitted in all expressions except F^h_0. The corrections δa, δb, δc, . . . are to be added to the values of a_0, b_0, c_0, . . . that are used in the F^h_0.

We are now ready to set up and solve the normal eqs. (13); also eqs. (12) if we wish to adjust the observations. The first normal equation will be

$$\sum_h \frac{x_h^2y_h^2}{a^2y_h^2/w_h+1/w_{n+h}}\,\delta a+\sum_h \frac{x_hy_h^2}{a^2y_h^2/w_h+1/w_{n+h}}\cdot\frac{\delta b}{b_0}+x_0y_0\lambda_{n+1}$$
$$=-\sum_h\frac{x_hy_h(b_0e^{a_0x_h}-y_h)}{a^2y_h^2/w_h+1/w_{n+h}}. \quad . \qquad (35)$$

This can be written as

$$\sum_h \frac{x_h^2}{a^2/w_h+1/u_h}\,\delta a+\sum_h \frac{x_h}{a^2/w_h+1/u_h}\,\delta\, ln\, b+x_0y_0\lambda_{n+1}$$
$$=-\sum_h\frac{x_h(ln b_0+a_0x_h-ln\, y_h)}{a^2/w_h+1/u_h}, \quad . \qquad (36)$$

wherein $u_h\equiv y_h^2w_{n+h}$.

This is exactly the same as the first normal equation that one would get by writing the empirical formula as

$$ln\, b+ax=ln\, y,$$

and considering this to be a linear equation in x and $ln\, y$, $ln\, y$ having a weight y^2 times the weight of y, which is, of course, in accordance with the law of propagation of error. In using the linear form of the exponential law, the change in weight that occurs when passing from y to $ln\, y$ is often ignored. The resulting erroneous procedure may yield values of a and b that are satisfactory, but it should not be confused with least squares.

No such pitfalls arise when the normal equations are formed by the procedure outlined above, for the result will be independent of the form in which the empirical equation is expressed. To illustrate this with the exponential law in the linear form we may write B for $ln\, b$ and f^h in place of F^h, where now

$$f^h(r_1, r_2, \ldots, r_{2n};\ a, B) = B + ar_h - ln\, r_{n+h}$$

and

$$\left.\begin{aligned} &f_0^h = B_0 + a_0 x_h - ln\, y_h, \\ &f_a^h = x_h,\ f_B^h = 1, \\ &f_h^h = a,\ f_{n+h}^h = -1/y_h. \\ &L_h = a^2/w_h + 1/y_h^2 w_{n+h} = a^2/w_h + 1/u_h. \end{aligned}\right\}.$$

u_h again denotes $y_h^2 w_{n+h}$. The first normal equation formed from these is clearly the same as eq. (36). The change in the mode of writing the empirical formula has produced no change in the results.

(c) The form $ya^x = b$.—The equation

$$ya^x = b \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

is closely allied with the exponential form, for by replacing a by $-ln\, a$ the equation $y = be^{ax}$ becomes $ya^x = b$. But since it is as easy to derive the normal equations anew as it is to make this substitution, we put

$$F^h(r_1, r_2, \ldots, r_{2n};\ a, b) = r_{n+h} a^{r_h} - b.$$

Approximate values of a and b can be found by plotting $\log y$ against x. Then

$$\left.\begin{aligned} &F_0^h = y_h a_0^{x_h} - b_0, \\ &F_a^h = x_h y_h a^{x_h - 1} \doteq x_h b/a,\ F_b^h = -1, \\ &F_h^h = y_h a^{x_h} ln\, a \doteq b\, ln\, a, \\ &F_{n+h}^h = a^{x_h} \doteq b/y_h, \\ &L_h = (b\, ln\, a)^2/w_h + a^{2x_h}/w_{n+h}. \end{aligned}\right\} \quad . \quad . \quad . \quad (38)$$

(d) *The indirect determination of e and h* *.—The determination of e and h* from ionization potentials, photoelectric effect, and other methods that do not measure e or h directly, is an important problem that is covered by the formulæ of the last section. e and h must satisfy

$$h=A_n e^{n/3}, \quad \ldots \ldots \ldots \quad (39)$$

wherein the A_n is determined experimentally †‡. n has the values 3, 4, 5, 6, which correspond to the different methods for determining h from e in eq. (39). If w_{A_n} is the weight of A_n, then $u_n \equiv w_{A_n}/e_0^{2n/3}$ is the weight of $h_n \equiv A_n e_0^{n/3}$. Birge's papers are largely concerned with the assignment of the weights u_3, u_4, u_5 from considerations of the experimental data.

We write

$$\left.\begin{aligned} &F^n \equiv h - A_n e^{n/3},\ n=3,\ 4,\ 5, \\ &F^n_0 = h_0 - A_n e_0^{n/3}, \\ &F^n_A = -e^{n/3}, \\ &F^n_e = -nh/3e,\ F^n_h = 1, \\ &L = e^{2n/3}/w_{A_n} = 1/u_n. \end{aligned}\right\} \quad \ldots \ldots \quad (40)$$

h/e may be taken as

$$6{\cdot}55\times10^{-27}/4{\cdot}77\times10^{-10}=1{\cdot}37\times10^{-17},$$

and h^2/e^2 as $1{\cdot}89\times10^{-34}$, in the last three lines and in the normal equations that follow.

In setting up the normal equations it may be well to make provision for forcing agreement with Eddington's theory, according to which the fine structure constant α is related to h and e by $h=2\pi e^2/\alpha c$ (e in e.s.u.). Since e enters here to the 6/3 power we introduce

$$-F^6_0 = -h_0 + (2\times137\pi/c)e_0^2 = -h_0 + 287{\cdot}1274\times10^{-10}e_0^2 \quad (41)$$

* Here e and h signify electron charge and Planck's constant, and, of course, have nothing to do with the e and h of the last section; and the n introduced in eq. (39) is distinct from the previous n. It will be convenient in this and the two following sections to use n in superscripts and subscripts where before we used h.

† Raymond T. Birge, Phys. Rev. Supplement (now Reviews of Modern Physics), i. pp. 1–73 (1929); Phys. Rev. xl. pp. 228–261 (1932).

‡ W. N. Bond, Phil. Mag. x. pp. 994–1003 (1930); *ibid.* xii. pp. 632–640 (1931).

in order to force α to be 137 exactly. A_6 may be used to denote the factor $287{\cdot}1274\times10^{-10}$.

The normal equations for determining e and h are easily formed by using the derivatives and weighting factors of eqs. (40) in eqs. (13). They are

$$\left.\begin{array}{llll} \delta e & \delta h & \lambda_6 & \text{Const.} \\ \hline \frac{h^2}{9e^2}(9u_3+16u_4+25u_5) & -\frac{h}{3e}(3u_3+4u_4+5u_5) & -2\frac{h}{e} & =\frac{h}{3e}(3u_3F_0^3+4u_4F_0^4+5u_5F_0^5) \\ & (u_3+u_4+u_5) & 1 & =-(u_3F_0^3+u_4F_0^4+u_5F_0^5) \\ & & 0 & =-F_0^6 \end{array}\right\}. \qquad (42)$$

The third row and column are to be deleted when agreement with Eddington's theory is to be released. The magnitude of the discrepancy between the two sets of values of e and h, obtained with and without the third row and column, is a basis for rejecting or accepting Eddington's theory.

Eqs. (42) show that the indirect determination of e and h is the same problem as fitting a line to the three points $(3, A_3e_0^{3/3})$, $(4, A_4e_0^{4/3})$, $(5, A_5e_0^{5/3})$ in the n, h plane, where the ordinates have weights u_3, u_4, u_5. This is just what Bond and Birge did. Bond noticed that with the correct value of e, individual determinations of h should not show a trend with the index n; but if they do, then the assumed value of e must be too high by the amount $3e/h$ times the slope of the line. Thus the slope of the line determines the adjusted value of e. Its intercept at $n=0$ is the adjusted value of h.

As an example, we may take Birge's "solution e" on p. 244 of his 1932 paper. He lists

n	$h_n\equiv A_n e_0^{n/3}$.
3	$(6{\cdot}5472\pm0{\cdot}0058)\times10^{-27}$
4	$(6{\cdot}5409\pm0{\cdot}0037)\times10^{-27}$
5	$(6{\cdot}5471\pm0{\cdot}0012)\times10^{-27}$
6	$6{\cdot}5330\times10^{-27}$

The weights u_3, u_4, u_5 are simply one of an infinite set of numbers that stand in the inverse ratios of the squares of the probable errors in h_n; *i. e.*,

$$u_3 : u_4 : u_5 = 1/0{\cdot}0058^2 : 1/0{\cdot}0037^2 : 1/0{\cdot}0012^2.$$

One possible set is u_3, u_4, $u_5 = 0{\cdot}30$, $0{\cdot}73$, $6{\cdot}94$, as used by Birge. Using $h_0 = 6{\cdot}547 \times 10^{-27}$ for forming the right-hand members of the normal equations, we obtain

$$\begin{array}{cccccl} 10^{10}\delta e & 10^{27}\delta h & 10^{27}\lambda_6 & & \text{Const.} & \\ \hline 39{\cdot}33 & -17{\cdot}62 & -2{\cdot}75 & = & 6{\cdot}48\times 10^{-3} & \\ & 7{\cdot}97 & 1 & = & -3{\cdot}70\times 10^{-3} & \\ & & 0 & = & -14{\cdot}00\times 10^{-3} & \end{array} \Bigg\} . \quad (43)$$

The solution is

$$10^{10}\delta e = 0{\cdot}0213,\ 10^{27}\delta h = 0{\cdot}0444,\ 10^{27}\ \lambda_6 = 0{\cdot}0174. \quad (44)$$

The adjusted values are

$$e = e_0 + \delta e = 4{\cdot}7913 \times 10^{-10},\ \ h = h_0 + \delta h = 6{\cdot}5914 \times 10^{-27}.$$

These give $hc/2\pi e^2 = 136{\cdot}999$, which proves the work, since agreement with Eddington's theory was forced.

By deleting the third row and column we, of course, obtain Birge's "solution e";

$$10^{10}\delta e = -0{\cdot}0048,\ 10^{-27}\delta h = -0{\cdot}0110,\ e = 4{\cdot}7652 \times 10^{-10},$$
$$h = 6{\cdot}5360 \times 10^{-27},\ hc/2\pi e^2 = 137{\cdot}400. \quad . \quad (45)$$

Because of the large discrepancy between the two pairs of e and h obtained with and without the third row and column, we should probably have to reject Eddington's theory on the basis of the data used for this illustration.

(e) *Direct measurement of* e *included.*—In the preceding section h and e entered eq. (39) as parameters. But if there are direct measurements on e also to be considered, e is no longer a parameter. To the list of observations there must be added e_d of weight w_e for the electronic charge. The empirical formula now is of the type $yz^x = b$. As we are using it, $x \equiv n/3$ and hence has infinite weight. The only changes come in the weighting factors. From eqs. (40) we now have

$$\left.\begin{array}{l} F_0^n = h_0 - A_n e_d^{n/3}, \\ L_{nm} = nmh^2/9e^2 w_e = nm/9u_e = L_{mn},\ m \neq n, \\ L_{nn} = e^{2n/3}/w_{An} + n^2h^2/9e^2 w_e = 1/u_n + n^2/9u_e. \end{array}\right\} . \quad (46)$$

Here $u_n \equiv w_{An}/e^{2n/3}$ as before, and $u_e \equiv w_e(e/h)^2$. u_e is accordingly the weight of the product of e_d by h_0/e_0 and can be compared with u_3, u_4, u_5, since it has the same dimensions.

Agreement with Eddington's theory is now forced by using $A_6 = 287 \cdot 1274 \times 10^{-10}$ with infinite weight. Then $1/u_6 = 0$ and

$$L_{66} = 36/9u_e. \quad \ldots \ldots \quad (47)$$

The normal equations come by inserting these into eqs. (7). They are

$$\begin{array}{cccccc}
\lambda_3 & \lambda_4 & \lambda_5 & \lambda_6 & \delta h & \text{Const.} \\
\hline
L_{33} & L_{34} & L_{35} & L_{36} & -1 & F_0^3 \\
 & L_{44} & L_{45} & L_{46} & -1 & F_0^4 \\
 & & L_{55} & L_{56} & -1 & F_0^5 \\
 & & & L_{66} & -1 & F_0^6 \\
 & & & & 0 & 0
\end{array}\Bigg\} . \quad . \quad (48)$$

δh comes directly from the normal equations, while by eq. (6)

$$\begin{aligned}
\delta e &= \frac{1}{3w_e}(3A_3\lambda_3 + 4A_4 e^{1/3}\lambda_4 + 5A_5 e^{2/3}\lambda_5 + 6A_6 e^{3/3}\lambda_6) \\
&= \frac{1}{3u_e}\frac{e}{h}(3\lambda_3 + 4\lambda_4 + 5\lambda_5 + 6\lambda_6) \\
&= (0 \cdot 243/u_e)(\lambda_4 + 2\lambda_5 + 3\lambda_6) \times 10^{17}. \quad \ldots \ldots \quad (49)
\end{aligned}$$

If agreement with Eddington's theory is not to be forced, we delete the fourth row and column; λ_6 does not then exist.

For an illustration we may take Birge's "solution l" *. Hereafter, no confusion will arise if for convenience the powers of 10 be omitted from the figures for h and e.

The data are

	With $e_0 = 4 \cdot 770$.	With $e_0 = 4 \cdot 768$.
$n=3$. .	$h_3 = 6 \cdot 5500 \pm 0 \cdot 0066$	$h_3 = 6 \cdot 5473 \pm 0 \cdot 0066$
$n=4$. .	$h_4 = 6 \cdot 5431 \pm 0 \cdot 0042$	$h_4 = 6 \cdot 5394 \pm 0 \cdot 0042$
$n=5$. .	$h_5 = 6 \cdot 5496 \pm 0 \cdot 0012$	$h_5 = 6 \cdot 5450 \pm 0 \cdot 0012$

* Raymond T. Birge, Phys. Rev. xl. p. 319 (1932). A numerical error in this was corrected by Birge at the Pasadena meeting of the American Physical Society, Dec. 16, 1932. See Phys. Rev. xliii. p. 211 (1933).

Millikan's direct measurement of e, $e_d=4{\cdot}768\pm0{\cdot}005$.

$F_0^n=h_0-h_n$, so by choosing $h_0=6{\cdot}5473$, we have

$$F_0^3=0,\ F_0^4=0{\cdot}0079,\ F_0^5=0{\cdot}0023,$$
$$F_0^6=h_0-287{\cdot}1274e_d^2=6{\cdot}5473-6{\cdot}5275=0{\cdot}0198.$$

The probable errors give the ratios

$$1/u_3 : 1/u_4 : 1/u_5 : 1/u_e=0{\cdot}0066^2 : 0{\cdot}0042^2 : 0{\cdot}0012^2 : (0{\cdot}005\times6{\cdot}55/4{\cdot}77)^2. \quad . \quad . \quad (50)$$

A convenient set is obtained by placing $1/u_e=1$, so we use

$$1/u_3,\ 1/u_4,\ 1/u_5,\ 1/u_e=0{\cdot}9241,\ 0{\cdot}3742,\ 0{\cdot}03055,\ 1.$$

$$\left.\begin{array}{llll} L_{33}=1{\cdot}92, & L_{34}=4/3, & L_{35}=5/3, & L_{36}=6/3, \\ & L_{44}=2{\cdot}15, & L_{45}=20/9, & L_{46}=24/9, \\ & & L_{55}=2{\cdot}81, & L_{56}=30/9, \\ & & & L_{66}=36/9. \end{array}\right\} \quad . \quad . \quad . \quad (51)$$

The normal equations are obtained at once by inserting these into eqs. (7).

λ_3	λ_4	λ_5	λ_6	δh		Const.
1·92	1·33	1·67	2	-1	=	0
	2·15	2·22	2·67	-1	=	$7{\cdot}9\times10^{-3}$
		2·81	3·33	-1	=	$2{\cdot}3\times10^{-3}$
			4	-1	=	$19{\cdot}8\times10^{-3}$
				0	=	0

. . . (52)

The solution is

$$\left.\begin{array}{l} \lambda_3=\ 14{\cdot}04,\ \lambda_4=23{\cdot}40,\ \lambda_5=-183{\cdot}84, \\ \lambda_6=146{\cdot}27,\ \delta h=43{\cdot}60,\ \text{all}\times10^{-3}. \end{array}\right\} \quad . \quad . \quad (53)$$

$$\delta e=0{\cdot}243(\lambda_4+2\lambda_5+3\lambda_6)=0{\cdot}0230 \quad . \quad . \quad . \quad . \quad (54)$$

The adjusted values are

$$\left.\begin{array}{r} e=4{\cdot}7680+0{\cdot}0230=4{\cdot}7910, \\ h=6{\cdot}5473+0{\cdot}0436=6{\cdot}5909, \\ hc/2\pi e^2=137{\cdot}006. \end{array}\right\} \quad . \quad . \quad (55)$$

With the systematic procedure previously recommended, δh, being written last, would be solved for first and its value automatically checked in the process. A short cut to finding e is to stop the solution at this point and use the adjusted value of $h=h_0+\delta h$ to compute $(hc/2\pi\times 137)^{\frac{1}{2}}$, which is necessarily the adjusted value of e, since we are forcing $hc/2\pi e^2$ to be 137.

By deleting the fourth row and column we release Eddington's theory and obtain

$$\left.\begin{array}{l}\lambda_3=-3{\cdot}931,\ \ \lambda_4=12{\cdot}661,\ \ \lambda_5=-8{\cdot}730,\ \ \delta h=-5{\cdot}288,\\ \delta e=(0{\cdot}243/u_e)\ (\lambda_4+2\lambda_5)=-1{\cdot}17,\ \ \text{all}\times 10^{-3}.\end{array}\right\} \quad (56)$$

Thence the adjusted values are

$$\left.\begin{array}{r}e=4{\cdot}7680-0{\cdot}0012=4{\cdot}7668,\\ h=6{\cdot}5473-0{\cdot}0053=6{\cdot}5420,\\ hc/2\pi e^2=137{\cdot}373.\end{array}\right\} \quad . \ . \ . \ . \quad (57)$$

These are, of course, identical with those that Birge presented at the Pasadena meeting.

Since these values of e and h are so different from those obtained by forcing Eddington's theory, one would probably conclude that, on the basis of the experimental data here used, there is little chance that $1/\alpha$ can be 137.

(*f*) *Direct measurements on both* e *and* h.—In section (*d*) neither e nor h was supposed to be directly observed; both therefore entered the condition equations $h=A_n e^{n/3}$ as parameters, the A_n being observed. In section (*e*), e and h must still satisfy the same conditions, but there are direct observations on e to be included, so that e is no longer a parameter. Both of these cases were successfully handled by Birge, following Bond's suggestion. Further than that, both Bond and Birge were able to estimate the probable errors of the adjusted values of e and h and of A_3, A_4, A_5, hence also of e/m, though their estimates were not in agreement. The presence of the foregoing two sections may serve as a means for understanding better the application of least squares in this problem.

There is a further reason for treating the determination of e and h in this paper, for by the methods here outlined it is a simple matter to handle the situation that will arise if some method is ever devised for finding h independently of e. Then neither e nor h will enter as parameters.

The empirical formula will be of the type $yz^x = w$, containing no parameters, so the normal equations will be in the form of eq. (8). Going back to eqs. (40) and (46) we see that the weighting factors off the diagonal will be unchanged, but those on the diagonal will annex a third term ; thus, now

$$\left.\begin{aligned} L_{nm} &= nmh^2/9e^2w_e = nm/9u_e = L_{mn}, \quad m \neq n, \\ L_{nn} &= e^{2n/3}/w_{\lambda n} + n^2h^2/9e^2w_e + 1/w_h = 1/u_n + n^2/9u_e + 1/w_h. \end{aligned}\right\} \quad (58)$$

Here w_h is the weight of the directly observed value of h. Its relation to the other weights would be found from the probable error in this determination by including it in the ratios of eq. (50), since w_h has the same dimensions as u_3, u_4, u_5, u_e.

The normal equations will be

$$\left.\begin{array}{ccccc} \lambda_3 & \lambda_4 & \lambda_5 & \lambda_6 & \text{Const.} \\ \hline L_{33} & L_{34} & L_{35} & L_{36} & = F_0^3 \\ & L_{44} & L_{45} & L_{46} & = F_0^4 \\ & & L_{55} & L_{56} & = F_0^5 \\ & & & L_{66} & = F_0^6 \end{array}\right\} \quad . \quad . \quad . \quad (59)$$

The approximate values of e and h to be used in computing F_0^n are those that are directly observed. The corrections to be added are

$$\delta e = \frac{1}{3u_e}\frac{e}{h}(3\lambda_3 + 4\lambda_4 + 5\lambda_5 + 6\lambda_6)$$

$$= (0{\cdot}243/u_e)\,(3\lambda_3 + 4\lambda_4 + 5\lambda_5 + 6\lambda_6) \times 10^{17}. \quad . \quad (60)$$

$$\delta h = -(1/w_h)\,(\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6). \quad . \quad . \quad . \quad . \quad . \quad (61)$$

The fourth row and column would, of course, be deleted if Eddington's theory is to be released.

The Assignment of Weights to Coordinates of different Dimensions.

Any empirical formula $F(x, y, z : a, b, c, \ldots) = 0$ may be considered as the equation of a surface on the orthogonal axes Ox, Oy, Oz. Values for the parameters in the formula can be determined from observational data, that is, from a set of observed points. To each

observed point near the surface there corresponds an adjusted point lying *on* the surface. As is evident from eqs. (10), the direction cosines δx_h, δy_h, δz_h of the line segment joining the observed and adjusted positions of any point will depend on the weights assigned to the coordinates at that point and on the graduation of the scales along the axes, whether the coordinates are dimensionally alike or unlike.

In order that the weighted squares of the residuals may be added together to form the quantity ϕ, the weight of x must have the dimensions $1/x^2$ and the weight of y the dimensions $1/y^2$. So if the x and y coordinates are different dimensionally, their weights will depend on the units in which x and y are expressed. The weights of the x coordinates cannot be assigned independently of the weights of the y coordinates; they must bear some relation to each other if the solution by least squares is to be unique. Only one of the qn weights can be assigned at will, and all the others must depend on it; for clearly if in

$$L_h = F^h_h F^h_h / w_h + F^h_{n+h} F^h_{n+h} / w_{n+h} + F^h_{2n+h} F^h_{2n+h} / w_{2n+h} \qquad (62)$$

the denominators bear no relation to each other, then each weighting factor L_h would be arbitrary and the normal equations would be arbitrary, and any solution whatever would be possible. Here, as in the first part of the paper, the subscripts h, $n+h$, $2n+h$ refer to x, y, z.

The question of how this relation between weights can be settled was raised in the first paper. It can now be answered by introducing the dependence of weights upon probable errors, and by requiring the different terms in the weighting factors L_h to be unequivocal save for a common constant multiplying factor. For by the elementary theory of least squares, the weight of any observation is inversely proportional to the square of the probable error of that observation; if, then, we write

$$w_i = C/t_i^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (63)$$

wherein t_i is the probable error of the ith observation, the constant of proportionality C must retain the same numerical value for all coordinates *regardless of their dimensions and units*; otherwise the terms in L_h would not bear a constant ratio to one another. The value

of C is arbitrary ; this is the one degree of freedom in the assignment of the weights.

If L_h is expressed in terms of the probable errors of the observations by means of eq. (63), it is clear from the law of propagation of error that L_h is $1/C$ times the square of the probable error in F^h arising from the probable errors of the x, y, z coordinates, and $1/L_h$ is the weight of F_h.

One of the qn weights is arbitrary, as has been stated. When any one of them is fixed, C is fixed, and this value of C must be used as the constant of proportionality for all the other weights. The task of assigning weights to the observations is thus seen to be identical with the task of estimating their probable errors.

In case there are two coordinates each of constant weight, and the data are extensive or have been repeated a number of times at each point, the evaluation of the relative weights of coordinates can be effected by a method published by W. R. Cook *.

When the probable errors of any function arising from the separate probable errors of the coordinates are known, it is advantageous to introduce auxiliary weights $u_1, u_2, u_3, \ldots$, defined by

$$\left.\begin{aligned} 1/u_h &\equiv F^h_h F^h_h/w_h = (1/C)\,(F^h_h t_h)^2, \\ 1/u_{n+h} &\equiv F^h_{n+h} F^h_{n+h}/w_{n+h} = (1/C)\,(F^h_{n+h} t_{n+h})^2, \\ 1/u_{2n+h} &\equiv F^h_{2n+h} F^h_{2n+h}/w_{2n+h} = (1/C)\,(F^h_{2n+h} t_{2n+h})^2. \end{aligned}\right\} \quad (64)$$

The right-hand sides are respectively $1/C$ times the squares of the probable errors in the function F^h arising from the probable errors in the x, y, and z coordinates, so under the conditions stipulated, the auxiliary weights are easily evaluated. This is the device that was used in the determination of e and h when u_n was introduced in place of $w_{An}/e^{2n/3}$ and u_e in place of $w_e(e/h)^2$. It was easier to work with u_n and u_e than with w_{An} and w_e, since the ratios $u_3 : u_4 : u_5 : u_e$ were readily evaluated from the published probable errors in h_3, h_4, h_5, e_d. The auxiliary weights are dimensionally similar and are, therefore, usually easy to evaluate.

The weighting factors then simplify to

$$L_h = 1/u_h + 1/u_{n+h} + 1/u_{2n+h}. \quad . \quad . \quad . \quad . \quad (65)$$

* W. R. Cook, Phil. Mag. xii. pp. 1025–1039 (1931).

The terms in L_h evidently influence the weight of the function F^h in the same manner that the masses of two bodies affect the "reduced mass" as used in mechanics for the orbit of one body revolving about another.

The Precision of the adjusted Values.

As emphasized in recent papers by Birge *, Shewhart †, and Eddington ‡, the "most probable" value of any quantity, computed by least squares or any other method, is of little value without a measure of its precision, so that there is a range of values associated with a given probability. It looks hopeful that there might be developed a systematic procedure for determining the probable errors of the adjusted values of the parameters and observations for the general case, since a preliminary survey § has shown that methods analogous to those now used in geodetic work can be elaborated so that the normal equations (13) will yield the probable errors of a, b, c, . . . as well as the corrections, the work being checked as it progresses. These results and further developments will appear in detail later. Henry Schultz ‖ and Birge ¶ have shown how to determine the probable error of a curve or surface and its extrapolation, when the probable errors of the parameters are known. Their papers also contain methods for finding the probable errors of the parameters when only one coordinate is subject to error.

In conclusion, it is a pleasure to thank Dr. H. F. Stimson and Dr. N. S. Osborne, of the United States Bureau of Standards, for their interest and counsel during the preparation of the manuscript. Several helpful suggestions were also made by Professor Raymond T. Birge of the University of California. The numerical calculations were made by Lola S. Deming.

* Raymond T. Birge, Phys. Rev. xl. pp. 207–227 (1932).

† W. A. Shewhart, "Probability as a Basis for Action," a paper read at a meeting of the American Mathematical Society, Atlantic City, December 28–30, 1933.

‡ A. S. Eddington, Proc. Phys. Soc. (London) xlv. pp. 271–287 (1933).

§ W. Edwards Deming, paper presented at the Chicago meeting of the American Physical Society, June 19, 1933. Abstract published in Phys. Rev. xliv. p. 317 (1933).

‖ Henry Schultz. J. Amer. Stat. Assoc. xxiv. Supplement, pp. 86–89 (March 1929); *ibid.* xxv. pp. 139–185 (1930).

¶ Raymond T. Birge, *loc. cit.*

LXVIII. *The Cathode Fall in the Hydrogen Arc.* By F. H. Newman, *D.Sc., F.Inst.P., Professor of Physics, University College of the South-West of England, Exeter* *.

THE cathode fall of potential in various metallic arcs has been measured by different experimenters, *e. g.*, Nottingham †, Compton and Lamar ‡, Killian §, Compton and Van Voorhis ||, Bramhall ¶, Dow, Attwood, and Timoshenko **, Langmuir ††, and Newman ‡‡. Such measurements are important because they indicate some of the processes occurring in the mechanism of the arc.

The distribution of potential when a current is flowing in the space between two electrodes, across which a potential difference is applied, assumes different forms in various cases, but in general most of the potential drop occurs within a very short distance—less than one millimetre—from the cathode, and in most cases this potential drop is approximately equal to the ionization potential of the metallic vapour through which the current passes. In addition, there are many types of electrical discharges and glows in which most of the potential fall occurs within a short distance from the cathode, the rest of the space having practically the potential of the anode. It is difficult, however, to correlate the value of the normal cathode fall in such electrical discharges with any constants of the cathode material, although as a rule the value increases with the work that has to be done to remove an electron from the cathode surface.

In the case of the cold cathode arc positive ions and electrons are generated throughout the space between the electrodes, but, if the rate of generation exceeds a certain value, there is a tendency to develop a potential maximum in the arc which is higher than that of both electrodes. As a result the low-velocity electrons produced by ionization are held in this region, and so tend to accumulate. This

* Communicated by the Author.
† Journ. Frank. Inst. ccvi. p. 43 (1928).
‡ Phys. Rev. xxxvii. p. 1069 (1931).
§ *Ibid.* xxxi. p. 1122 (1928).
|| Nat. Acad. Sci. Proc. xiii. p. 336 (1927).
¶ Phil. Mag. xiii. p. 682 (1932).
** Paper presented at the Summer Convention of the A.I.E.E., Chicago, June 1933.
†† Report No. 7, Section I., International Electrical Congress, Paris, 1932.
‡‡ Phil. Mag. xv. p. 601 (1933).

accumulation produces a negative space charge, and the potential of this region falls until it attains a value at which the electrons just escape to the anode—the initial velocities of the electrons enabling them to move against the small retarding field. The region in which a potential maximum develops has been called by Langmuir a *plasma*, and is a relatively field-free space—the concentration of the electrons being approximately equal to that of the positive ions. In addition, the recombination of the ions and electrons in a plasma takes place at a practically negligible rate, since the electrons move in orbits among the ions and rarely collide with them.

If a collector is introduced into the arc it becomes covered with an almost non-luminous *sheath*, the thickness of which increases with the potential difference between the collector and the space in the arc at which the collector is placed. The region of the arc disturbed by the collector is strictly limited, the flow of ions to the collector being controlled by the space-charge which they excite in its neighbourhood—the sheath acting as an electrostatic shield to the remainder of the gas. If the collector potential is negative with respect to the space in which it is situated, a sheath of positive ions will form round the collector and there will be reflexion of the electrons which, while penetrating the sheath, have insufficient energy to pass to the collector against the field set up by the maintained potential on the collector. On the other hand, positive ions which penetrate the sheath pass on to the collector, and the rate at which these ions strike it is a measure of the current in the circuit containing the collector. This positive ion current does not change much, as the collector potential is increased with respect to the cathode, until it reaches a value slightly less than that of the space surrounding the collector when fast-moving electrons penetrate the sheath and, striking the collector, give rise to an appreciable electron-current. The total current in the collector circuit is then the sum of the two—electron and positive ion currents.

When the collector potential is made greater than that corresponding to equality of electron and positive ion currents, the electron-current exceeds the latter, and the logarithm of the electron-current value i_e varies linearly with the potential v of the collector until the

potential of the latter is raised to a point where it becomes equal to that of the plasma, and the positive ion sheath disappears. If the collector potential is raised above this point, however, i tends to remain constant, for the collector is receiving all the electrons which are flowing toward it in the plasma. Raising the potential above this point merely repels the positive ions in the plasma and an electron-sheath develops over the collector, but the potential cannot be greatly increased without accelerating the electrons to such an extent that further ionization occurs, the electron-sheath disappearing and larger currents flowing to the collector. There is, however, a change in slope of the log $i_e - v$ curve where the collector is at the same potential as the space, and the position of this change of slope determines the space-potential.

When the collector is a fine wire, and is used to measure the potential in an electric arc, the bombardment of the collector by positive ions is so violent that it either melts or is raised to such a high temperature that it emits electrons, this electronic emission rendering the results invalid. It has been found, however, that a suitable collector in such a case can be made by turning a tungsten rod to a diameter of one millimetre over a length of 10 mm. at one end. The latter is then covered with a thin glass tube and the extremity ground flat. The remaining part of the tungsten rod is 6 mm. diameter. Such a collector was used with the sodium arc, and has now been employed to measure the cathode fall in a hydrogen arc.

In previous papers * the author has described what may be termed a cold cathode arc, such an arc being started and maintained between cold tungsten electrodes. The radiation emitted is characteristic of the gases in the tube, but there are no tungsten lines. In the present experiments hydrogen gas was used with specially purified tungsten electrodes, placed 5 mm. apart in a cylindrical glass discharge-tube. The hydrogen was purified and dried, the gas-pressure in the tube being 10^{-2} mm. of mercury, as at this pressure the arc could be satisfactorily started and maintained—the ignition being effected by passing a momentary electrical discharge between one of the arc electrodes and a third one in the tube.

The steady potential difference applied across the arc electrodes was 200 volts, with a ballast resistance in the

* Phil. Mag. xiv. p. 788 (1932).

external circuit to regulate the current. The anode and cathode potential drops, with the arc passing, were measured by means of the collector method described above, but the arc was stopped as soon as the anode became red-hot; the cathode always remained comparatively cool.

In practice various currents ranging from 5 to 10 amperes were employed, but both the cathode and anode falls were constant, so far as could be detected, over this range. The values were: cathode fall 28·5 volts, anode fall 7·4 volts, the centre of the collector being 2 mm. from the electrode in each case. The total potential difference between the two electrodes varied from 39 volts to 37 volts over the current range employed.

These values obtained for the electrode falls of potential are obviously too high, because the collector cannot be placed sufficiently close to the electrode. It has been calculated by Compton and Lamar * that the distance over which the cathode fall of potential extends in a mercury arc is less than $1{\cdot}76 \times 10^{-4}$ cm.

Other experiments indicate that the value of the cathode fall in an arc of metallic vapour is in the neighbourhood of the ionizing potential of the vapour: for example, the cathode fall for a tungsten arc is 16·2 volts †, whereas Anderson and Kretchmar ‡ found that the total potential difference across a tungsten arc was 14 volts at 12 amperes.

It is obvious in the present case, however, that the ionizing potential of tungsten does not enter into the problem, because with pure iron electrodes practically identical results for the two electrode potential drops were obtained. On the other hand, the cathode fall is greater than the ionization potential of hydrogen, viz., 16·1 volts, although the probability of ionization taking place at an electronic collision, when it is energetically possible for it to occur, increases rapidly above the ionization potential. In addition, the type of arc described is intermediate in character between a purely metallic arc and an electrical discharge. In the latter the normal cathode fall, although varying with the metal, gas, and gas-pressure, is several hundred volts in value.

* *Loc. cit.*
† Phys. Rev. xxi. p. 266 (1923).
‡ *Ibid.* xxvi. p. 33 (1925).

LXIX. *On the Resonance Frequency of Oscillatory Circuits with Leaky Condenser and its Bearing on the Measurement of the Dielectric Constant of Ionized Gas.* By

S. S. Banerjee, *M.Sc., Demonstrator in Physics, Hindu University, Benares* *.

Introduction.

THE theory of the bending of wireless waves by the ionized upper atmosphere originated by Eccles and Larmor [5] has, in recent years, stimulated experimental investigations [1], [2], [3], [4] to provide a direct proof of the reduction of the dielectric constant of a medium by the presence of free ions and electrons in it.

The principle of the experimental method usually employed consists in measuring the change in the capacity of a condenser when the space between its plates is filled either with an ionized gas in which both electrons and gas molecules are present or with pure electronic atmosphere as may be found in a highly evacuated thermionic valve.

Results of investigation in both these cases show anomalous change in the value of the dielectric constant. The change is found to be sometimes positive, sometimes negative, and sometimes nil, that is, the dielectric constant of the medium between the plates is sometimes found to be greater, sometimes less, and sometimes equal to unity, though the simple theory of Eccles and Larmor indicates that the dielectric constant should always be less than unity. The anomalous results obtained in experiments with pure electronic atmosphere has been discussed by B. C. Sil in a paper recently published in this Magazine [9]. The anomalous results in the case of ionized gas are considered in this paper.

In experiments with ionized gas the gas is contained in a discharge-tube and the condenser plates are placed sometimes inside and sometimes outside the tube. In the former case the apparent increase has been attributed to the formation of ionic sheaths round the surface of the metallic condenser plates which has the effect of reducing the distance between them thereby increasing its capacity. In the latter case, that of condenser plates outside the discharge-tube, Gutton and Clement [1] have tried to

* Communicated by Prof. S. K. Mitra, D.Sc.

explain the increase of dielectric constant recorded by them as due to the existence of quasi-elastically bound electrons in the ionized gas. Pedersen[8], however, is inclined to attribute the increase to the effect of the conductivity acquired by the ionized gas inside the discharge-tube. That the conductivity will have the apparent effect of increasing the dielectric constant has been recognized by Appleton and his co-workers. To eliminate this conductivity effect they have taken the precaution of measuring the resonance condition of the circuit by noting the setting of the measuring condenser for maximum voltage across its plates—the voltage resonance—instead of the more usual one for maximum current in the circuit—the current resonance. But in spite of such precautions experimenters have recorded anomalous variations of the dielectric constant of an ionized gas.

We shall show in the present paper that the conductivity of the ionized gas is the real cause of the apparent increase of the dielectric constant and that the dispositions of condenser system and the discharge-tube in the usual experiments are such that even measurement by voltage resonance does not eliminate the effect of the conductivity. A significant fact about the results obtained by research workers in this field is that the increased value in dielectric constant is noticed when the ionization is strong, *i. e.*, when the conductivity is large, it being remembered that in the Eccles–Larmor theory[5] there is no room for recording an increase in dielectric constant at any ionic concentration.

2. *Theoretical Considerations.*

The resonance frequency of an oscillatory circuit can be regarded from two standpoints; firstly, the frequency for which the current flowing through the whole circuit is a maximum ("current resonance") and, secondly, the frequency for which the voltage across any condenser or a set of condensers in the system is a maximum ("voltage resonance"). In the case of a simple circuit consisting of a condenser and an inductance the two resonant frequencies, as obtained by noting the current or the voltage resonance, have the same value. If the dielectric between the plates of the condenser possesses

conductivity then the resonant frequencies as obtained by the two methods are no longer the same.

We shall now discuss how the resonant frequencies are affected, firstly, when the conductivity of the space between the plates is uniform and, secondly, when it is non-uniform.

Case I.—*Conducting material uniformly distributed.*

When the conducting material is uniformly distributed throughout the space between the plates, fig. 1 (*a*), the system behaves like a condenser shunted by a resistance as shown in fig. 1 (*b*). The condition for voltage resonance is obtained by determining the condition for which the voltage developed across the condenser C is a maximum when the circuit is loosely coupled to an oscillatory circuit of frequency ω.

Fig. 1 (*a*).

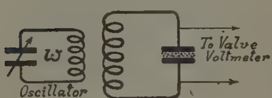
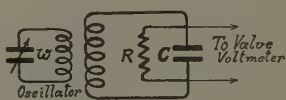


Fig. 1 (*b*).



The instantaneous current flowing in the circuit is given by

$$\frac{e}{j\omega L+\dfrac{1}{\dfrac{1}{R}+j\omega C}},$$

where e is the instantaneous voltage acting in the circuit and L, C, and R have their usual meanings.

The voltage across the condenser C is given by

$$\frac{e\left(\dfrac{1}{\dfrac{1}{R}+j\omega C}\right)}{j\omega L+\dfrac{1}{\dfrac{1}{R}+j\omega C}}=\frac{e}{\dfrac{j\omega L}{R}-\omega^2 CL+1}.$$

The condition for maximum voltage is obtained by equating the denominator to zero.

This gives

$$\frac{j\omega L}{R}=0 \quad \text{and} \quad 1-\omega^2 CL=0,$$

or

$$\omega=\frac{1}{\sqrt{CL}}.$$

This is the well-known condition for current resonance in a circuit consisting of a capacity C and an inductance L. Thus when the conductivity in the space between the condenser plates is uniform the resonant frequency, as measured by noting the condition for maximum voltage across C, is independent of the conductivity and is equal to the resonant frequency when the space is a non-conductor.

Case II.—*Conducting material non-uniformly distributed.*

The conductivity of the medium is represented by the dotted portion of the condenser C, fig. 2 (*a*). For

Fig. 2 (*a*).

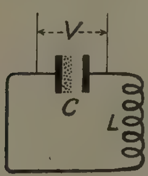
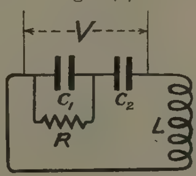


Fig. 2 (*b*).



simplification of calculation we assume that the dotted portion is uniformly conducting whereas the clear portion is non-conducting.

The space between the condenser C as a whole is therefore non-uniformly conducting. The condenser is equivalent to the condenser system C_1 and C_2 $\left(\frac{1}{C}=\frac{1}{C_1}+\frac{1}{C_2}\right)$ of fig. 2 (*b*), the resistance R representing the inverse of conductivity of the dotted portion of C, the capacity of which is C_1.

With the same notations as in case I. the instantaneous current in the circuit is

$$\frac{e}{j\omega L+\frac{1}{j\omega C_2}+\frac{1}{\frac{1}{R}+j\omega C_1}}.$$

The instantaneous voltage V, across the set of condensers C_1 and C_2 in fig. 2 (*b*), is given by

$$\frac{e\left(\dfrac{1}{\dfrac{1}{R}+j\omega C_1}+\dfrac{1}{j\omega C_2}\right)}{j\omega L+\dfrac{1}{j\omega C_2}+\dfrac{1}{\dfrac{1}{R}+j\omega C_1}}=\frac{e}{\dfrac{j\omega L}{\dfrac{1}{\dfrac{1}{R}+j\omega C_1}+\dfrac{1}{j\omega C_2}}+1}.$$

The denominator of the above expression can be written as

$$\frac{j\omega C_2R+1+j\omega C_1R-\omega^2LC_2-j\omega^3LC_1C_2R}{1+j\omega C_1R+j\omega C_2R}.$$

Rationalizing the denominator of this expression, we get

$$\frac{\begin{aligned}1+\omega^2R^2(C_1+C_2)^2-\omega^2LC_2-j\omega^3LC_1C_2R\\+j\omega^3LC_2R(C_1+C_2)-\omega^4LC_1C_2R^2(C_1+C_2)\end{aligned}}{1+\omega^2R^2(C_1+C_2)^2}.$$

Now, the voltage across the set of condensers C_1 and C_2, fig. 2 (*b*), will be maximum when the above expression is minimum.

Equating the real part of the above expression to zero, we get

$$\omega^4LC_1C_2R^2(C_1+C_2)+\omega^2\{LC_2-R^2(C_1+C_2)^2\}-1=0.$$

Solving the above equation we get

$$\omega^2=\frac{-\{LC_2-R^2(C_1+C_2)^2\}\pm\sqrt{\{LC_2-R^2(C_1+C_2)^2+4R^2LC_1C_2(C_1+C_2)}}{2LC_1C_2R^2(C_1+C_2)}.$$

Neglecting the negative and imaginary values of ω, we get for its real positive value

$$\omega=\sqrt{\frac{-\{LC_2-R^2(C_1+C_2)^2\}+\sqrt{\{LC_2-R^2(C_1+C_2)^2+4R^2LC_1C_2(C_1+C_2)}}{2LC_1C_2R^2(C_1+C_2)}}.$$

Thus in such arrangement the resonant frequency is *not* independent of the conductivity.

A closer inspection of the above formula shows that ω increases with the increase of R. In fact, it may be readily deduced that, when $R=0$,

$$\omega=\frac{1}{\sqrt{LC_2}},$$

and when $R=\infty$,

$$\omega=\frac{1}{\sqrt{L\frac{C_1C_2}{C_1+C_2}}}.$$

It may be mentioned here that if the voltmeter in fig. 2(*b*) be connected across the shunted condenser C_1 only, then the resonant frequency is independent of R. This has been shown by A. Astin (6).

3. *Experiments and Results.*

The principle of the experimental arrangement for determining the change in dielectric constant produced

Fig. 3 (*a*).

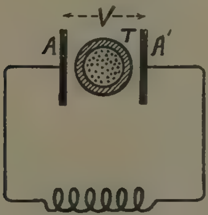
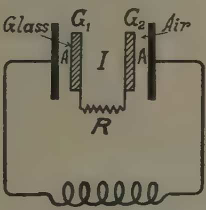


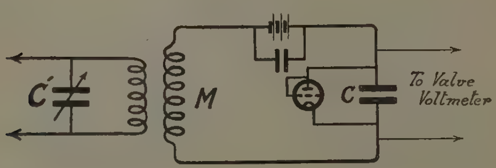
Fig. 3 (*b*).



by ionization is shown in fig. 3(*a*). AA′ are the condenser plates and T the cross-section of the discharge-tube. The resonant frequency of the oscillatory circuit is measured first when the gas is in normal state and again when it is ionized by passing electric discharge through it. The resonant frequency is obtained by noting the frequency at which the voltage across the condenser plates is a maximum. Now the condenser system of the circuit is obviously equivalent to that shown in fig. 3 (*b*). An air condenser A, a glass condenser G_1, an air condenser I (shunted by a resistance R representing the inverse of conductivity of the ionized gas), a glass condenser G_2, and an air condenser A are all in series. This condenser

system again is, in effect, equivalent to the condenser system shown in fig. 2(*b*), namely, two condensers C_1 and C_2 in series, one of which is shunted by a resistance R. Now it has been shown in the previous section that in such a condenser system the voltage resonance, that is, the frequency for which the voltage across the condenser system attains a maximum value, is not independent of the value of the resistance R. It is easy to see, without going through the calculations of the previous section, that the resonant frequency will decrease with the decrease of the value of the resistance. If R is zero the resonant frequency will simply be due to a circuit of inductance L and capacity C_2. If, on the other hand, R is infinite the resonant frequency will be that due to a circuit of inductance L and capacity $\frac{C_1C_2}{C_1+C_2}$. In order to check the results obtained in the previous section the experiments described below were carried out.

Fig. 4.



(*a*) Experiment to prove that in the circuit fig. 1 (*b*) the voltage resonance *is independent* of the shunting resistance R.

It is difficult to obtain a conducting substance which is free from self-inductance or self-capacity. The only conductor that we could think of was that due to the plate-filament space of a valve rendered conducting by heating the filament. The experimental arrangement is shown in fig. 4.

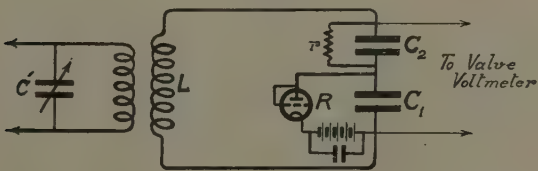
The coil of the valve oscillator is loosely coupled to that of the measuring circuit M. Resonance is obtained by varying the frequency of the oscillator by varying the value of the condenser C^1 and noting the minimum deflexion of the thermionic voltmeter connected across C. Reading is taken once when the filament is cold and again

when the filament is hot. It will be noticed that since the valve is present during both the observations the effect of its capacity is merely to increase the total capacity of the circuit M. Careful observation showed that the conductivity of the plate-filament space had only the effect of altering the sharpness of the resonance curve. The position of the peak of the voltage resonance curve was not affected in any way.

The experiment was also repeated with salt solution, by dipping the two plates of the condenser in the salt solution and changing the concentration. As expected, the resonant frequency was not found to change.

An interesting point noted in this connexion was that increase of the plate voltage had the effect of sharpening the resonance showing, as if the resistance of the valve had increased. This is due to the fact that the increased voltage increases the speed of the electrons which are

Fig. 5.



thus made to remain in the plate-filament space of the valve for a shorter time. Each electron carrying the current from filament to plate is subjected to the alternating field for only a small fraction of the complete cycle. The effect is discussed in a paper by B. C. Sil [9], recently published in the Phil. Mag., and also by Benner [7] in a paper in *Ann. der Phys.*.

(*b*) Experiments to prove that in circuit of fig. 2(*b*) the voltage resonanace frequency *is not independent* of the resistance.

The experimental arrangement is as shown in fig. 5.

The notations are the same as in fig. 2 (*b*). R represents the differential resistance of the valve at the particular voltage of the anode.

In this case, in order to complete the anode circuit of the valve, a high resistance *r* was put across C_2. as shown in

the figure. The drop of voltage due to this resistance was made up by increasing the value of H.T. Battery. The value of r was much greater than $\omega^2 C_2^2$ and also than R, the differential valve resistance. The shunting effect of this resistance across C_2 could thus be neglected.

Everything being connected up as shown in the diagram, resonant frequency was observed once when the filament was cold and again when it was hot. The frequency was observed to change. Observations were taken at three different frequencies, which are given in the table. The resonant frequencies were also calculated with the help of the formula given in last section. It will be seen that the observed frequencies agree fairly closely with the calculated frequencies.

The table shows that the (voltage) resonance frequency is lowered when the plate-filament space of the valve R,

L.	C_1.	C_2.	R.	Resonance frequency with filament cold. Observed.	Resonance frequency with filament hot. Observed.	Calculated.
39·6	68·6	539	1866	3370	2804	2879
91·8	49·6	539	1866	2190	1948	2000
231·3	706	539	1866	618	585	598

L, C_1, C_2, and R refer to fig. 5. L in microhenry, C_1 and C_2 in micro-microfarad, R in ohm, and frequency in kilocycle/sec.

shunting the condenser C_1, is rendered conducting by electrons emitted from its heated filament.

The experiment was also repeated with salt solution, by putting the salt solution in a rectangular glass cell outside which the two plates of a condenser were kept in close contact and observing the resonance with different concentrations. A change (decrease) was observed in the resonant frequency when the conductivity was increased.

4. *Summary.*

In the experimental investigation on the change of dielectric constant of an ionized gas according to Eccles–Larmor Theory, the assumption is usually made that

the voltage resonance frequency of the experimental oscillatory circuit is independent of the conductivity of the ionized gas which fills part of the space between the plates of the oscillatory circuit condenser. This assumption is studied critically, and it is shown theoretically and experimentally that the voltage resonance frequency is independent of the conductivity only when the medium between the condenser plates is uniformly conducting. If the conductivity is not uniform the voltage resonance frequency is not independent of the conductivity. The bearing of these results on the measurement of the depression of the dielectric constant of an ionized gas is discussed. It is shown that the anomalous results heretofore reported by various investigators can be explained as due to the conductivity acquired by the gas owing to ionization. The dispositions of the condenser system and the discharge-tube in these experiments are such that measurement of voltage resonance does not eliminate the effect of conductivity produced by ionization.

The present investigation was carried out under the direction of Prof. S. K. Mitra, D.Sc., in the Wireless Laboratory of the University College of Science, Calcutta, during the summer vacation of 1933. I wish to tender my sincere thanks for his keen interest and constant help throughout the course of the investigation.

I further wish to express my thanks to Prof. P. Dutt, M.A. (Cantab), Head of the Physics Department, Hindu University, Benares, for his kindness in granting me facility to carry out the investigation at Calcutta.

References.

(1) Gutton and Clement, *Comptes Rendus*, clxxxiv. (1927).
(2) Gutton, *Annals de Physique*, tome xiii. p. 68 (1930).
(3) Appleton and Childs, Phil. Mag. ser. 7, x. p. 989 (1930).
(4) Appleton and Chapman, Proc. Phys. Soc. xliv. p. 246 (1932).
(5) Larmor, Phil. Mag. (6) xlviii. no. 228, p. 1025 (1924).
(6) A. Astin, Phys. Rev. xxxiii. p. 1074 (1929).
(7) Benner, *Ann. der Phys.* ser. 5, Band 3, p. 993 (1928).
(8) Pedersen, The Propagation of Radio Waves (Danmarks Naturvidenskabelige Samfund), page 94.
(9) B. C. Sil, Phil. Mag. (7) xvi. p. 1114 (1933),

Physics Laboratory,
Hindu University,
Benares, India.

LXX. *Note on the Dynamical Theory of Commutator Motors.* By W. H. Ingram, *Columbia University, New York* †.

THE equations of armature e.m.f. for an orthocyclic alternating-current slip-ring machine obtained by Park ‡ have a form

$$\left.\begin{aligned}\sqrt{2}e_d&=\tfrac{2}{3}r\dot{\xi}+\dot{\phi}_d-\dot{\theta}\phi_t,\\ \sqrt{2}e_t&=\tfrac{2}{3}r\dot{\eta}+\dot{\phi}_t+\dot{\theta}\phi_d,\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (1)$$

which is identical with the voltage-equations of the commutator machine of fig. 1. This is made clear § by a consideration of the voltage-drops in the conductor shown in fig. 3 where we must have, for the time-interval during which the circuit is complete, the relation

$$\mathrm{V}(t)=\begin{pmatrix}\text{resistance-}\\ \text{drop}\end{pmatrix}+\frac{d\phi_t}{dt}+\begin{pmatrix}\text{velocity of}\\ \text{conductor}\end{pmatrix}$$
$$\times\begin{pmatrix}\text{conductor-}\\ \text{length}\end{pmatrix}\times\begin{pmatrix}\text{longitudinal component}\\ \text{of field-strength}\end{pmatrix},$$

where ϕ_t is the transverse component of the flux and where the last term is equal to $\dot{\theta}\times$*flux*.

[The numerical coefficients and algebraic signs arise from a choice of definition and convention which has seemed most convenient and natural in the theory of slip-ring motors and generators ‖. The steady-state equations for the salient-pole alternator,

$$\left.\begin{aligned}e_d&=r\dot{q}_d-x_t\dot{q}_t,\\ e_t&=r\dot{q}_t+x_d\dot{q}_d+e_0{}^*,\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

may be obtained from (1) with the simplifying substitution

$$\left.\begin{aligned}\tfrac{1}{3}\sqrt{2}\dot{\xi}&=\dot{q}_d=\dot{q}_0\sin(\gamma-\psi),\\ \tfrac{1}{3}\sqrt{2}\dot{\eta}&=\dot{q}_t=\dot{q}_0\cos(\gamma-\psi),\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (3)$$

and are diagrammed in fig. 5.]

† Communicated by the Author.

‡ Park, Trans. Am. Inst. Elect. Eng. xlviii. p. 716 (1929).

§ The resolution of the e.m.f. at the brushes of a commutator machine into a transformer voltage, a d.-c. generator voltage, and an ohmic drop was made in the classical paper by Atkinson on the theory of the commutator machine (Proc. Inst. Civil Eng. cxxxiii. p. 113, 1898); and earlier by Potier and Görges, *vide* Dyhr, 'Die Einphasen Motoren,' Berlin, 1912.

‖ Ingram, Phil. Mag. xvii. p. 271 (1934).

When the brushes are inclined, as in fig. 2, at an angle β, we may write directly

$$\left.\begin{aligned}\sqrt{2}e_\xi &= \tfrac{2}{3}r\dot{\xi}+\dot{\phi}_\xi-\dot{\theta}\phi_\eta,\\ \sqrt{2}e_\eta &= \tfrac{2}{3}r\dot{\eta}+\dot{\phi}_\eta+\dot{\theta}\phi_\xi,\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (4)$$

where ϕ_ξ is the flux in the direction of the ξ-axis, inclined at an angle β with the direct axis, and where ϕ_η is the flux in the direction of the η-axis.

When there are an infinite number of commutator segments, and when adjacent rotor conductors are infinitely close, the flux which surrounds one of the rotor conductors, due to the current in it, also surrounds its neighbour. At the instant when the current in a conductor is broken at the commutator an exactly equal

Fig. 1.

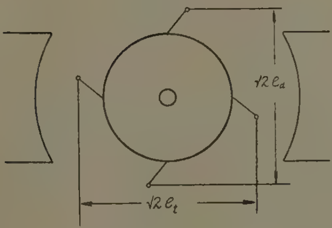
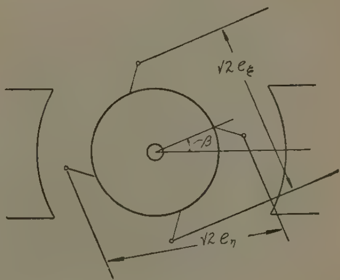


Fig. 2.



current is established in its neighbour. Hence the theory of a commutator machine with an infinite number of commutator segments does not have to take account of brush-width and local brush-currents.

The fluxes are linear functions of the currents, and the equations †

$$\left.\begin{aligned}\phi_d &= (\tfrac{2}{3}l^*+l_\mu)\dot{\xi}_d+m_0\cos\alpha\, i+\mathrm{M}_0\mathrm{I},\\ \phi_t &= (\tfrac{2}{3}l^*-l_\mu)\dot{\eta}_t+m_0\sin\alpha\, i,\end{aligned}\right\} \quad . \quad . \quad (5)$$

where i is the amortisseur current and I the field-current, define the coefficients as inductances. Here, clearly, $\dot{\xi}_d$ is the component of the current which produces flux in the direct axis and $\dot{\eta}_t$ a component which produces flux

† A term is to be added on the right to the second of this pair when there is an inter-polar field winding.

in the transverse axis. The fluxes in the directions of the axes inclined at the angle β are given by

$$\left.\begin{aligned}\phi_\xi&=\phi_d\cos\beta-\phi_t\sin\beta,\\ \phi_\eta&=\phi_d\sin\beta+\phi_t\cos\beta,\end{aligned}\right\}\quad . \quad . \quad . \quad . \quad . \qquad (6)$$

in accordance with fig. 4. The currents are related by a similar diagram and by the equations

$$\left.\begin{aligned}\dot\xi&=\dot\xi_d\cos\beta-\dot\eta_t\sin\beta,\\ \dot\eta&=\dot\xi_d\sin\beta+\dot\eta_t\cos\beta,\end{aligned}\right\}\quad . \quad . \quad . \quad . \quad . \qquad (7)$$

Fig. 3.

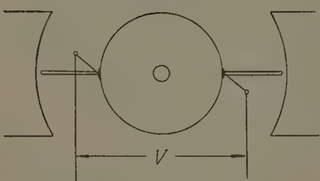
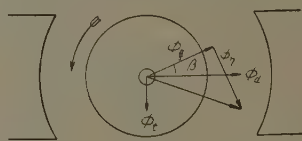


Fig. 4.



The substitution of $\dot\xi_d$ and $\dot\eta_t$, from (7), into (5) and of ϕ_d and ϕ_t, from (5), into (6) yields

$$\left.\begin{aligned}\phi_\xi=(\tfrac{2}{3}l^*+l_\mu\cos 2\beta)\dot\xi+l_\mu\sin 2\beta\,\dot\eta&\\ +m_0\cos(\alpha+\beta)i+\mathrm{M}_0\mathrm{I}\cos\beta,&\\ \phi_\eta=(\tfrac{2}{3}l^*-l_\mu\cos 2\beta)\dot\eta+l_\mu\sin 2\beta\,\dot\xi&\\ +m_0\sin(\alpha+\beta)i+\mathrm{M}_0\mathrm{I}\sin\beta.&\end{aligned}\right\}\quad . \qquad (8)$$

When the brushes are rotating, the equations for the brush-voltages become

$$\left.\begin{aligned}\sqrt{2}e_\xi&=\tfrac{2}{3}r\dot\xi+\dot\phi_\xi-(\dot\theta-\dot\beta)\phi_\eta,\\ \sqrt{2}e_\eta&=\tfrac{2}{3}r\dot\eta+\dot\phi_\eta+(\dot\theta-\dot\beta)\phi_\xi.\end{aligned}\right\}\quad . \quad . \quad . \qquad (9)$$

This follows from the fact that the conductor in contact with the brushes has a speed $(\dot\theta-\dot\beta)$ relative to the moving flux component which it cuts.

The equations (9) may be obtained dynamically in a way similar to the way equations (1) were obtained †.

† Ingram, *loc. cit.*

Consider the equations of motion of the three-phase star-connected alternating-current motor :

$$\left.\begin{aligned} e_k &= r\dot{q}_k + \frac{d}{dt}[\Sigma l_{jk}\dot{q}_j + m_k i + \mathrm{M}_k \mathrm{I}], \\ 0 &= \mathbf{R}i + \frac{d}{dt}[\Sigma m_k \dot{q}_k + \mathbf{L}i + \mathbf{M}\mathrm{I}], \\ \mathrm{E} &= \mathrm{RI} + \frac{d}{dt}[\Sigma \mathrm{M}_k \dot{q}_k + \mathbf{M}i + \mathrm{LI}], \\ -f &= \Theta\ddot{\theta} + s\dot{\theta} - \frac{\partial \mathrm{T}}{\partial \theta}, \end{aligned}\right\} \quad . \quad . \quad . \quad (10)$$

where $\qquad l_{jk} = l^* 1_{jk} - l_m - l_\mu \cos(2\theta - (j+k)2\pi/3),$

where 1_{jk} is a quantity equal to unity when $j=k$ and otherwise zero.

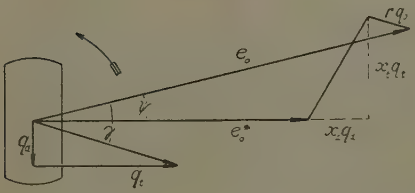


Fig. 5.

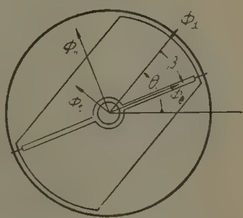


Fig. 6.

When the variables $\dot{\xi}, \dot{\eta}$ are connected with the $\dot{q}$'s by the equations

$$\left.\begin{aligned} \dot{\xi} &= \Sigma\dot{q}_i \sin(\theta_i - \beta), \\ \dot{\eta} &= \Sigma\dot{q}_i \cos(\theta_i - \beta), \\ \dot{\zeta} &= \Sigma\dot{q}_i, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (11)$$

the first equation of (10) may be transformed to (9) on making use of the identities

$$\Sigma l_{jk}\dot{q}_j = l^*\dot{q}_k - l_m\dot{\zeta} + l_\mu\dot{\xi}\sin(\theta_k + \beta) - l_\mu\dot{\eta}\cos(\theta_k + \beta),$$

$$\Sigma m_k\dot{q}_k = \dot{\xi}m_0\cos(\alpha + \beta) + \dot{\eta}m_0\sin(\alpha + \beta),$$

$$\Sigma \mathrm{M}_k\dot{q}_k = \dot{\xi}\mathrm{M}_0\cos\beta + \dot{\eta}\mathrm{M}_0\sin\beta,$$

$$\Sigma \mathrm{M}_k'\dot{q}_k = \dot{\xi}\mathrm{M}_0\sin\beta + \dot{\eta}\mathrm{M}_0\cos\beta,$$

$$\sqrt{2}e_\xi = \tfrac{2}{3}\Sigma e_k \sin(\theta_k - \beta).$$

$$\sqrt{2}e_\eta = \tfrac{2}{3}\Sigma e_k \cos(\theta_k - \beta).$$

A machine with stationary armature is represented in fig. 6. The terms in (9) proportional to $(\dot\theta-\dot\beta)$, the velocity of the brushes, may be called the Coriolis voltage components, because of their analogy to the forces of Coriolis in dynamics †. Thus, noting that the kinetic energy, in terms of the new coordinates, is

$$\begin{aligned}T=&\tfrac{1}{2}(\tfrac{2}{3}l^*+l_\mu\cos 2\beta)\dot\xi^2+\tfrac{1}{2}(\tfrac{2}{3}l^*-l_\mu\cos 2\beta)\dot\eta^2+l_\mu\sin 2\beta\,\dot\xi\dot\eta\\&+m_0\cos(\alpha+\beta)\,\dot\xi i+m_0\sin(\alpha+\beta)\,\dot\eta i+\tfrac{1}{2}\mathbf{L}i^2\\&+\mathrm{M}_0\cos\beta\,\dot\xi\mathrm{I}+\mathrm{M}_0\sin\beta\,\dot\eta\mathrm{I}+\tfrac{1}{2}\mathrm{LI}^2\\&+\tfrac{1}{2}(\tfrac{1}{3}l^*-l_m)\dot\zeta^2+\mathbf{M}i\mathrm{I}+\tfrac{1}{2}\Theta\dot\theta^2\quad .\ .\ .\ .\ .\ .\ (12)\end{aligned}$$

and paraphrasing a paragraph in Whittaker (§ 88, 3rd ed.) where

$(\dot\xi, \dot\eta, \dot\zeta)$ is to be written for (u, v, w),
(e_d, e_t) ,, ,, ,, (X, Y),

it is found immediately that the equations (8) are obtained when the angular velocity of the brushes $(\dot\theta-\dot\beta)$ with respect to the armature conductors is taken as the angular velocity of the axes of reference. The equations of motion are given by the system

$$\left.\begin{aligned}\frac{\partial(\mathrm{U}-\mathrm{S})}{\partial\dot\xi}&=\frac{d}{dt}\left(\frac{\partial\mathrm{T}}{\partial\dot\xi}\right)-(\dot\theta-\dot\beta)\frac{\partial\mathrm{T}}{\partial\dot\eta},\\\frac{\partial(\mathrm{U}-\mathrm{S})}{\partial\dot\eta}&=\frac{d}{dt}\left(\frac{\partial\mathrm{T}}{\partial\dot\eta}\right)+(\dot\theta-\dot\beta)\frac{\partial\mathrm{T}}{\partial\dot\xi},\\\frac{\partial(\mathrm{U}-\mathrm{S})}{\partial i}&=\frac{d}{dt}\left(\frac{\partial\mathrm{T}}{\partial i}\right),\\\frac{\partial(\mathrm{U}-\mathrm{S})}{\partial\mathrm{I}}&=\frac{d}{dt}\left(\frac{\partial\mathrm{T}}{\partial\mathrm{I}}\right),\\\frac{\partial(\mathrm{U}-\mathrm{S})}{\partial\dot\theta}&=\frac{d}{dt}\left(\frac{\partial\mathrm{T}}{\partial\dot\theta}\right)-\frac{\partial\mathrm{T}}{\partial\theta},\end{aligned}\right\}\quad .\ .\ .\ (13)$$

where, it is to be noted, $\partial\dot\xi/\partial\theta=\dot\eta$, $\partial\dot\eta/\partial\theta=-\dot\xi$.

An expression similar to (12) may be obtained without difficulty when the brush-pairs are not perpendicular to each other (as in the Déri motor). When a brush is lifted from the commutator, the brush-current is written

† *Cf.* Whittaker, p. 217.

equal to zero (an equation of constraint), and when a brush-pair is short-circuited the brush-voltage is written equal to zero.

The system of equations (15) is simpler than its equivalent in terms of tensors, and no less general. Systems of machines are treated by adding the functions of specification (kinetic energy, potential energy, dissipation, and activity function) of the component machines, and the connexions of the system are displayed by the variables employed in these functions and by auxiliary equations of constraint.

LXXI. *Heat Conduction in a Solid in Contact with a well-stirred Liquid.* By ARNOLD N. LOWAN *.

A SOLID, in the form of a cylindrical bar, with arbitrary cross-section, extends from $x=0$ to $x=a$, the face $x=0$, and the lateral surface being impervious to heat. The solid is in contact with a well-stirred liquid, contained in a reservoir of the same geometrical shape as the solid, and extending from $x=a$ to $x=b$. The boundary $x=b$ is in contact with a medium of constant temperature, the heat transfer between the reservoir and the medium being proportional to the temperature difference, the coefficient of proportionality being designated by σ. What is the thermal history of the solid?

This problem has been treated by Schumann † and Langer ‡ by methods which are an extension of the Fourier analysis, and by March and Weaver § by means of the theory of the Volterra integral equation of the second kind. A considerably simpler solution may be obtained by the method of the Laplace transformation, as presented in two other papers ‖ ¶, in conjunction with the theory of loaded integral equations.

* Communicated by the Author.

† Phys. Rev. xxxvii. (1931).

‡ Tohoku, Math. J. xxxv. pp. 260–275 (1932).

§ Phys. Rev. June 1928, pp. 769–775.

‖ Arnold N. Lowan, "On the Cooling of a Radioactive Sphere." Phys. Rev., Nov. 1, 1933. This paper will subsequently be referred to as A. N. L.

¶ Arnold N. Lowan, "On the Problem of the Heat Recuperator," Phil. Mag. xvii. [in the press].

The problem reduces to the determination of a function $T(x, t)$, satisfying the following equations:

$$\frac{\partial T}{\partial t} - k\frac{\partial^2 T}{\partial x^2} = 0, \qquad t>0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\lim_{t\to 0} T(x, t) = f(x), \qquad 0<x<a, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

$$\left.\begin{array}{l}\lim_{t\to 0} T(a, t) = \theta = \text{initial temperature of liquid} \\ \qquad\qquad \text{in reservoir,}\end{array}\right\} \quad (3)$$

$$\frac{\partial T}{\partial x} = 0, \qquad x=0, \quad . \quad . \quad (4)$$

$$P\frac{\partial T}{\partial t} + K\frac{\partial T}{\partial x} - \sigma(T_1 - T) = 0, \qquad x=a, \quad . \quad . \quad (5)$$

where, in addition to the symbols previously defined, we have: P=heat capacity of a prism of liquid of length b and unit cross-section, and T_1=constant temperature of the medium.

Equation (5) may be written in the form

$$A\frac{\partial T}{\partial t} + \frac{\partial T}{\partial x} - C(T_1 - T) = 0, \qquad x=a, \quad . \quad (5')$$

where $A=\frac{P}{K}$ and $C=\frac{\sigma}{K}$. To solve the system (1) to (5), let us make the substitution

$$T(x, t) = \phi(x) + u(x, t), \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where $\phi(x)$ is as yet an undetermined function. Then the function $u(x, t)$ must satisfy the equations

$$\frac{\partial u}{\partial t} - k\frac{\partial^2 u}{\partial x^2} = k\frac{\partial^2 \phi}{\partial x^2}, \qquad t>0, \quad . \quad . \quad . \quad (7)$$

$$\lim_{t\to 0} u(x, t) = f(x) - \phi(x), \qquad 0<x<a, \quad . \quad (8)$$

$$\lim_{t\to 0} u(a, t) = \theta - \phi(a), \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

$$\frac{\partial u}{\partial x} + \phi'(x) = 0, \qquad x=0; \quad . \quad (10)$$

$$A\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} + Cu + C\phi(a) + \phi'(a) - cT_1 = 0, \quad x=a. \quad . \quad (11)$$

If we subject the system (7) and (11) to the Laplace transformation, then the function $y(x, \lambda) = \mathrm{L}\{u(x, t)\}$ * must satisfy the equations

$$ky'' - \lambda y = -\{f(x) - \phi(x)\} - \frac{k}{\lambda}\phi''(x), \quad . \quad . \quad (12)$$

$$y'(0) + \phi'(0) = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

$$\mathrm{A}\{\lambda y(a) + \phi(a) - \theta\} + y'(a) + \mathrm{C}y(a) + \frac{1}{\lambda}\{\phi'(a) + \mathrm{C}\phi(a) - c\mathrm{T}_1\} = 0. \quad . \quad . \quad . \quad (14)$$

If we determine the function $\phi(x)$ by the conditions

$$\phi(a) - \theta = \phi'(0) = 0, \quad . \quad . \quad . \quad . \quad (15\text{–}16)$$

$$\phi'(a) + \mathrm{C}\phi(a) - c\mathrm{T}_1 = 0, \quad . \quad . \quad . \quad . \quad . \quad (17)$$

the system (12) to (14) assumes the simpler form

$$ky'' - \lambda y = -\{f(x) - \phi(x)\} - \frac{k}{\lambda}\phi''(x), \quad . \quad . \quad (18)$$

$$y'(0) = (\mathrm{C} + \mathrm{A}\lambda)y(a) + y'(a) = 0. \quad . \quad (19\text{–}20)$$

The system (18), (19), and (20) differs from the system (10), (3), (4′) of A. N. L., in that the boundary condition (20) contains the parameter λ. For this reason the characteristic functions corresponding to the homogeneous differential equation

$$ky'' - \lambda y = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

in conjunction with the boundary conditions (19) and (20), no longer form a set of orthogonal functions. Instead, they satisfy the condition

$$\overline{\int_0^a} y_m(\xi) y_n(\xi)\, d\xi \equiv \int_0^a y_m(\xi) y_n(\xi)\, d\xi + \mathrm{A}k y_m(a) y_n(a) = 0, \quad \text{if } m \neq n, \quad (22)$$

as can be readily ascertained †.

If we define the appropriate Green function $\mathrm{G}(x, \xi, \lambda)$ as the solution of (21), (19), and (20), satisfying the usual

* See A. N. L.

† A. Kneser, ‘Integralgleichungen,’ p. 118 (1922). Kneser refers to an equation of type (22) as the “loaded” (Belastete) orthogonality condition.

discontinuity condition, then it can be shown in the usual manner*, that

$$y(x, \lambda)=\int_0^a \mathrm{G}(x, \xi, \lambda)\left\{f(\xi)-\phi(\xi)+\frac{k}{\lambda}\phi''(\xi)\right\}d\xi \quad (23)$$

is the solution of (18), (19), and (20).

From (23) we obtain, in the usual manner,

$$u(x, t)=\int_0^a \{f(\xi)-\phi(\xi)\}.\mathrm{L}^{-1}\{\mathrm{G}(x, \xi, \lambda)\}d\xi$$
$$+k\int_0^a \phi''(\xi).\mathrm{L}^{-1}\left\{\frac{1}{\lambda}\mathrm{G}(x, \xi, \lambda)\right\}d\xi,$$

or, with the aid of (9) A. N. L., and since $\mathrm{L}^{-1}\left\{\frac{1}{\lambda}\right\}=1$,

$$u(x, t)=\int_0^a \{f(\xi)-\phi(\xi)\}.\Gamma(x, \xi, t)\,d\xi$$
$$+k\int_0^a \phi''(\xi)\,d\xi.\int_0^t \Gamma(x, \xi, t-\tau)\,d\tau, \quad (24)$$

where

$$\Gamma(x, \xi, t)=\mathrm{L}^{-1}\{\mathrm{G}(x, \xi, \lambda)\}.$$

Kneser has shown that, for the "loaded" boundary problem of the type under consideration, the bilinear formula

$$\mathrm{G}(x, \xi, \lambda)=\sum_{n=1}^{\infty}\frac{y_n(x)y_n(\xi)}{\lambda_n+\lambda} \quad . \quad . \quad . \quad . \quad (25)$$

is valid, provided the characteristic functions satisfy the "loaded" normalizing condition

$$\overline{\int_0^a}\{y_n(\xi)\}^2 d\xi \equiv \int_0^a \{y_n(\xi)\}^2 d\xi+\mathrm{A}k\{y_n(a)\}^2=1. \quad (26)$$

In view of (24) and (25), our solution $\mathrm{T}(x, t)$ ultimately becomes

$$\mathrm{T}(x, t)=\phi(x)+\sum_{n=1}^{\infty} e^{-k\lambda_n^2 t}y_n(x)\int_0^a f(\xi)y_n(\xi)\,d\xi$$
$$-\sum_{n=1}^{\infty} e^{-k\lambda_n^2 t}y_n(x)\int_0^a \phi''(\xi)y_n(\xi)\,d\xi$$

* A. Kneser, *loc. cit.* vi. pp. 120–22. See also J. Teichmann, "Mechanische Problems die auf belastete Integralgleichungen fuhren," Dissertation. (Breslau, 1919.)

$$+k\sum_{n=1}^{\infty}\frac{y_n(x)}{k\lambda_n^2}\int_0^a \phi''(\xi)y_n(\xi)\,d\xi$$

$$-k\sum_{n=1}^{\infty}\frac{e^{-k\lambda_n^2 t}}{k\lambda_n^2}\int_0^a \phi''(\xi)y_n(\xi)\,d\xi, \quad . \quad . \quad (27)$$

where the summation extends over the characteristic values and the corresponding "loaded" normalized characteristic functions of the system

$$y''+\lambda_n^2 y=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

$$y'(0)=(\mathrm{C}-\mathrm{A}k\lambda_n^2)y_n(a)+y'(a)=0, \quad . \quad (29\text{–}30)$$

obtained from (21), (19), and (20) by replacing λ by $-k\lambda_n^2$.

Our solution (27) contains the undetermined function $\phi(x)$, which we now proceed to eliminate. Let

$$\mathrm{E}_n=\int_0^a \phi''(\xi)y_n(\xi)\,d\xi. \quad . \quad . \quad . \quad . \quad (31)$$

In view of (15), (16), (17), (28), (29), and (30), E_n may ultimately be written in the two alternative forms

$$\mathrm{E}_n=(\mathrm{CT}_1-\mathrm{A}k\lambda_n^2\theta)\,.\,y_n(a)-\lambda_n^2\int_0^a \phi(\xi)y_n(\xi)\,d\xi, \quad . \quad (32)$$

$$\mathrm{E}_n=\mathrm{CT}_1 y_n(a)-\lambda_n^2\overline{\int_0^a}\phi(\xi)y_n(\xi)\,d\xi, \quad . \quad . \quad . \quad . \quad (32')$$

the "loaded" integral in the last equation being defined as in (22) and (26).

Substituting for E_n from (32) and (32′), in the third and fourth term of (27), and making use of the identities

$$\sum_{n=1}^{\infty}\frac{y_n(x)y_n(a)}{k\lambda_n^2}=\Big\{\mathrm{G}(x,\xi,\lambda)\Big\}_{\substack{\xi=a\\ \lambda=0}}\equiv\frac{1}{\mathrm{C}} \quad . \quad . \quad (33)$$

(this can readily be verified by actually constructing the Green function for the case $\lambda=0$, and putting $\xi=a$), and *

$$\phi(x)=\sum_{n=1}^{\infty}y_n(x)\overline{\int_0^a}\phi(\xi)y_n(\xi)\,d\xi, \quad . \quad . \quad . \quad (34)$$

* A. Kneser, *loc. cit.* p. 121.

our solution (27) becomes

$$T(x,t)=T_1+\sum_{n=1}^{\infty} e^{-k\lambda_n^2 t}$$

$$\cdot y_n(x)\left\{\int_0^a f(\xi)y_n(\xi)\,d\xi-\frac{\sigma T_1-kP\lambda_n^2\theta}{K\lambda_n^2}\,y_n(a)\right\}, \quad (35)$$

where we have replaced A and C by $\frac{P}{K}$ and $\frac{\sigma}{K}$, respectively.

The characteristic values are readily seen to be the positive roots of the transcendental equation

$$(C-Ak\lambda^2)\cos a\lambda-\lambda\sin a\lambda=0. \quad . \quad . \quad . \quad (36)$$

Also the characteristic functions are

$$y_n(x)=\frac{1}{N_n}\cos\lambda_n x, \quad . \quad . \quad . \quad . \quad . \quad (37)$$

where, in view of (22), we have

$$N_n^2=\frac{a}{2}+\frac{1}{4\lambda_n}\sin 2\lambda_n a+\frac{P}{K}\cos^2\lambda_n a. \quad . \quad . \quad (38)$$

With this significance of N_n our solution (35) finally becomes

$$T(x,t)=T_1+\sum_{n=1}^{\infty}\frac{1}{N_n^2}\cdot e^{-k\lambda_n^2 t}$$

$$\cdot\cos\lambda_n x\left\{\int_0^a f(\xi)\cdot\cos\lambda_n\xi\,d\xi-\frac{\sigma T_1-kP\lambda_n^2\theta}{K\lambda_n^2}\cdot\cos\lambda_n a\right\}, \quad . \quad . \quad . \quad (39)$$

the summation being extended over the positive roots of (36).

With the aid of (34), it may be shown that (39) satisfies the equations (1) to (5), provided that $f(x)$ is piecewise continuous and twice differentiable. It thus represents the complete solution of our problem.

LXXII. *A Rapid Method for the Summation of a Two-Dimensional Fourier Series.* *By* C. A. Beevers *and* H. Lipson, *The George Holt Physics Laboratory, The University of Liverpool* *.

1. *Introduction.*

In some recent work on the structure of copper sulphate pentahydrate the authors had occasion to use the method of two-dimensional Fourier synthesis described by W. L. Bragg †. Copper sulphate is triclinic and possesses only the centro-symmetry which makes the Fourier method applicable, so that it represents the most general case which can arise. Thus all reflexions, such as $hk0$ and $hk0$, must be considered separately, and also the projection of full half of the unit-cell must be obtained. In crystals of higher symmetry both of these factors are often considerably reduced. For these reasons it was feared that the work on copper sulphate would be very long, and an attempt was therefore made to devise a short method of calculation.

That such a method is possible suggested itself in the following manner. The terms which have to be added are

$$F_{hk0} \cos 2\pi(hx+ky).$$

Terms with the same k may be regarded as producing, for a particular value of x, sinusoidal distributions of density with the same wave-length but different amplitudes and phase differences. The sum of these is a sinusoidal curve of calculable amplitude and phase angle. Thus the possibility arises of grouping together planes with the same value of k, and in this way reducing tremendously the size of the final tables. However, it does not seem possible to devise a rapid method of taking phase angle into account as such, but by the simple expansion of $\cos 2\pi(hx+ky)$ an equivalent process can be carried out.

2. *The New Method.*

The summation to be effected may be written as

$$\sum_{-k}^{k} \sum_{0}^{h} F_{hk0} \cos 2\pi(hx+ky). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

* Communicated by Professor L. R. Wilberforce, M.A.

† Proc. Roy. Soc. A, cxxiii. p. 537 (1929).

$$= \sum_{-k}^{k} \sum_{0}^{h} F_{hk0}\{\cos 2\pi hx . \cos 2\pi ky - \sin 2\pi hx . \sin 2\pi ky\}$$

$$= \sum_{0}^{h} \left[\sum_{-k}^{k} F_{hk0} \cos 2\pi ky \right] \cos 2\pi hx$$

$$- \sum_{0}^{h} \left[\sum_{-k}^{k} F_{hk0} \sin 2\pi ky \right] \sin 2\pi hx.$$

The factors in the brackets may be evaluated first. Routine methods can easily be devised for this. If we put

$$\sum_{-k}^{k} F_{hk0} \cos 2\pi ky = C_y, \quad \text{and} \quad \sum_{-k}^{k} F_{hk0} \sin 2\pi ky = S_y,$$

the expression becomes

$$\sum_{0}^{h} C_y \cos 2\pi hx - \sum_{0}^{h} S_y \sin 2\pi hx. \quad . \quad . \quad . \quad (2)$$

C_y and S_y are independent of x, and thus the values of the two components of (2) may be found for a particular value of y and every value of x. Here again, of course, routine methods are used.

3. *Details of the Method.*

Either the summations C_y and S_y may be made first or the corresponding summations C_x and S_x. It is best to sum with respect to that coordinate which makes the groups larger and less numerous, as in this way the final summation which has to be carried out over the whole area will be made as short as possible. For example, in $CuSO_4.5H_2O$ the measured planes were 89 in number and had values of h between ± 8 and k between ± 13. Thus C_y and S_y would have a maximum of 27 planes in each group, but the number of different groups would be only 9, corresponding to the values of h from 0 to 8. The final tables thus contain 9 terms instead of 89.

The evaluation of C_y and S_y for the different values of h is a small part of the work, but even this may be speeded up, since the $hk0$ and $h\bar{k}0$ reflexions add together in the cosine term and subtract in the sine term. Table I. shows the arrangement of the calculations of

$$\sum_{-k}^{k} F_{3k0} \cos 2\pi ky \quad \text{and} \quad \sum_{-k}^{k} F_{3k0} \sin 2\pi ky.$$

Table I.—Evalutions of $\sum_{-k}^{k} F_{3k0} \cos 2\pi ky$ and $\sum_{-k}^{k} F_{3k0} \sin 2\pi ky$.

$F_{3k0} \cos 2\pi ky$.

			y																									
h.	k.	F.	·00	·02	·04	·06	·08	·10	·12	·14	·16	·18	·20	·22	·24	·26	·28	·30	·32	·34	·36	·38	·40	·42	·44	·46	·48	·50
3	±1	40	40	40	39	37	35	32	29	26	22	18	12	8	2	$\bar{2}$	$\bar{8}$	$\bar{12}$	$\bar{18}$	$\bar{22}$	$\bar{26}$	$\bar{29}$	$\bar{32}$	$\bar{35}$	$\bar{37}$	$\bar{39}$	$\bar{40}$	$\bar{40}$
3	±2	$\bar{28}$	$\bar{28}$	$\bar{27}$	$\bar{25}$	$\bar{21}$	$\bar{16}$	$\bar{8}$	$\bar{2}$	5	13	18	23	26	28	28	26	23	18	13	5	$\bar{2}$	$\bar{8}$	$\bar{16}$	$\bar{21}$	$\bar{25}$	$\bar{27}$	$\bar{28}$
3	±3	44	44	41	32	19	3	$\bar{13}$	$\bar{28}$	$\bar{39}$	$\bar{44}$	$\bar{43}$	$\bar{35}$	$\bar{24}$	$\bar{9}$	9	24	35	43	44	39	28	13	$\bar{3}$	$\bar{19}$	$\bar{32}$	$\bar{41}$	$\bar{44}$
3	±4	18	18	16	10	1	$\bar{7}$	$\bar{15}$	$\bar{18}$	$\bar{17}$	$\bar{12}$	3	5	13	18	18	13	5	3	$\bar{12}$	$\bar{17}$	$\bar{18}$	$\bar{15}$	$\bar{7}$	1	10	16	18
3	$\bar{5}$	12	12	10	4	$\bar{4}$	$\bar{10}$	$\bar{12}$	$\bar{10}$	$\bar{4}$	4	10	12	10	4	$\bar{4}$	$\bar{10}$	$\bar{12}$	$\bar{10}$	$\bar{4}$	4	10	12	10	4	$\bar{4}$	$\bar{10}$	$\bar{12}$
3	±7	38	38	24	$\bar{8}$	$\bar{34}$	$\bar{36}$	$\bar{11}$	21	38	27	$\bar{2}$	$\bar{30}$	$\bar{37}$	$\bar{17}$	17	37	30	2	$\bar{27}$	$\bar{38}$	$\bar{21}$	11	36	34	8	$\bar{24}$	$\bar{38}$
3	9	7	7	3	$\bar{4}$	$\bar{7}$	$\bar{2}$	5	6	0	$\bar{6}$	$\bar{5}$	2	7	4	$\bar{4}$	$\bar{7}$	$\bar{2}$	5	6	0	$\bar{6}$	$\bar{5}$	2	7	4	$\bar{3}$	$\bar{7}$
3	±11	19	19	4	$\bar{18}$	$\bar{10}$	14	15	$\bar{9}$	$\bar{18}$	1	19	6	$\bar{17}$	$\bar{12}$	12	17	$\bar{6}$	$\bar{19}$	$\bar{1}$	18	9	$\bar{15}$	$\bar{14}$	10	18	$\bar{4}$	$\bar{19}$
Total	..	..	150	111	30	$\bar{19}$	$\bar{19}$	$\bar{7}$	$\bar{11}$	$\bar{9}$	5	18	$\bar{5}$	$\bar{14}$	18	74	92	61	24	$\bar{3}$	$\bar{15}$	$\bar{29}$	$\bar{39}$	$\bar{27}$	$\bar{21}$	$\bar{60}$	$\bar{133}$	$\bar{170}$

$F_{3k0} \sin 2\pi ky$.

3	±1	$\bar{2}$	0	0	0	$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{2}$	$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$	0	0	0
3	±2	14	0	3	7	10	12	13	14	14	13	11	8	5	2	$\bar{2}$	$\bar{5}$	$\bar{8}$	$\bar{11}$	$\bar{13}$	$\bar{14}$	$\bar{14}$	$\bar{13}$	$\bar{12}$	$\bar{10}$	$\bar{7}$	$\bar{3}$	0
3	±3	$\bar{8}$	0	$\bar{3}$	$\bar{5}$	$\bar{7}$	$\bar{8}$	$\bar{8}$	$\bar{6}$	$\bar{4}$	$\bar{1}$	2	5	7	8	8	7	5	2	$\bar{1}$	$\bar{4}$	$\bar{6}$	$\bar{8}$	$\bar{8}$	$\bar{7}$	$\bar{5}$	$\bar{3}$	0
3	$\bar{5}$	12	0	$\bar{7}$	$\bar{12}$	$\bar{12}$	$\bar{7}$	0	7	12	12	7	0	$\bar{7}$	$\bar{12}$	$\bar{12}$	$\bar{7}$	0	7	12	12	7	0	$\bar{7}$	$\bar{12}$	$\bar{12}$	$\bar{7}$	0
3	9	7	0	6	5	$\bar{2}$	$\bar{7}$	$\bar{4}$	4	7	3	$\bar{4}$	$\bar{7}$	$\bar{1}$	6	6	$\bar{1}$	$\bar{7}$	$\bar{4}$	3	7	4	$\bar{4}$	$\bar{7}$	$\bar{2}$	5	6	0
	±11	$\bar{5}$	0	$\bar{5}$	$\bar{2}$	4	3	$\bar{3}$	$\bar{5}$	1	5	1	$\bar{5}$	$\bar{2}$	4	4	$\bar{2}$	$\bar{5}$	1	5	1	$\bar{5}$	$\bar{3}$	3	4	$\bar{2}$	$\bar{5}$	0
Total	..	..	0	$\bar{6}$	$\bar{7}$	$\bar{8}$	$\bar{8}$	$\bar{3}$	13	28	30	15	$\bar{1}$	0	6	2	$\bar{10}$	$\bar{17}$	$\bar{7}$	4	0	$\bar{15}$	$\bar{29}$	$\bar{32}$	$\bar{28}$	$\bar{21}$	$\bar{12}$	0

Table II.

Final Summations for $y = \cdot 12$.

$C_y \cos 2\pi hx$.

h.	x. ·00	·04	·08	·12	·16	·20	·24	·28	·32	·36	·40	·44	·48
0..	28	28	28	28	28	28	28	28	28	28	28	28	28
1..	$\overline{30}$	$\overline{29}$	$\overline{27}$	$\overline{22}$	$\overline{16}$	$\bar{9}$	$\bar{2}$	6	13	19	24	28	30
2..	21	19	11	1	$\bar{9}$	$\overline{17}$	$\overline{21}$	$\overline{19}$	$\overline{17}$	$\bar{4}$	7	15	20
3..	$\overline{11}$	$\bar{8}$	$\bar{1}$	7	11	9	2	$\bar{6}$	$\overline{11}$	$\overline{10}$	$\bar{3}$	5	10
4..	$\bar{1}$	..	..	..	..	..	..	..	..	..	..	..	..
5..	$\overline{19}$	$\bar{6}$	15	15	$\bar{6}$	$\overline{19}$	$\bar{6}$	15	15	$\bar{6}$	$\overline{19}$	$\bar{6}$	15
6..	$\bar{8}$	0	8	2	$\bar{8}$	$\bar{2}$	7	3	$\bar{7}$	$\bar{4}$	6	5	$\bar{6}$
7..	18	$\bar{3}$	$\overline{17}$	10	13	$\overline{15}$	$\bar{8}$	17	1	$\overline{18}$	6	16	$\overline{12}$
8..	8	$\bar{3}$	$\bar{5}$	8	$\bar{2}$	$\bar{6}$	7	0	$\bar{7}$	6	2	$\bar{8}$	4
$\frac{1}{2}F_{000}$..	65	65	65	65	65	65	65	65	65	65	65	65	65
Total..	71	63	77	114	76	34	72	109	80	76	116	148	154

$S_y \sin 2\pi hx$.

1..	0	$\bar{8}$	$\overline{16}$	$\overline{22}$	$\overline{28}$	$\overline{30}$	$\overline{33}$	$\overline{32}$	$\overline{30}$	$\overline{25}$	$\overline{19}$	$\overline{12}$	$\bar{4}$
2..	0	47	82	98	88	58	13	$\overline{36}$	$\overline{75}$	$\overline{96}$	$\overline{90}$	$\overline{67}$	$\overline{25}$
3..	0	9	13	10	2	$\bar{8}$	$\overline{13}$	$\overline{11}$	$\bar{3}$	6	12	12	5
4..	0	6	6	1	$\bar{5}$	$\bar{6}$	$\bar{2}$	5	7	3	$\bar{4}$	$\bar{7}$	$\bar{3}$
5..	0	$\overline{17}$	$\overline{11}$	11	17	0	$\overline{17}$	$\overline{11}$	11	17	0	$\overline{17}$	$\overline{11}$
6..	0	7	1	$\bar{7}$	$\bar{2}$	6	3	$\bar{6}$	$\bar{3}$	6	4	$\bar{5}$	$\bar{5}$
7..	0	5	$\bar{2}$	$\bar{4}$	3	3	$\bar{4}$	$\bar{1}$	5	$\bar{1}$	$\bar{5}$	2	4
Total..	0	49	73	87	75	23	$\overline{53}$	$\overline{92}$	$\overline{88}$	$\overline{90}$	$\overline{102}$	$\overline{94}$	$\overline{39}$
Difference.	71	14	4	27	1	11	125	201	168	166	218	242	193
Sum	71	112	150	201	151	57	19	17	$\bar{8}$	$\overline{14}$	14	54	115
	1·00	·96	·92	·88	·84	·80	·76	·72	·68	·64	·60	·56	·52
	x.												

In cases where $3k0$ and $3\bar{k}0$ are combined the fact is indicated by the signs $\pm$ over the k index. It will be noticed that each line is symmetrical about the line $y=\cdot25$, and this may be used either to lighten the work or to serve as a check.

Table II. shows an example of the final summation for $y=\cdot12$. The range of calculation need extend only from $x=0$ to $x=\cdot50$, since both $\cos 2\pi hx$ and $\sin 2\pi hx$ are symmetrical about $x=\cdot50$, the latter with a change of sign. Thus from $x=0$ to $x=\cdot50$ the required total is the difference between the two summations, and from $x=\cdot50$ to $x=1\cdot00$ the sum.

As has already been pointed out, the calculations have been carried out for the most general type of crystal, and therefore cover any difficulties which may arise. With crystals of higher symmetry the work involved may be still further reduced. For example, if the crystal has an orthogonal lattice one of the terms in the expansion of $\cos 2\pi(hx+ky)$ will vanish, owing to the equality in magnitude of F_{hk0} and $F_{h\bar{k}0}$.

4. *Summary.*

By expanding $\cos 2\pi(hx+ky)$ the longest double Fourier synthesis can be accomplished by two workers in about two days. Tables are given which illustrate the method of procedure.

LXXIII. *Notices respecting New Books.*

Die Methoden zur angenäherten Lösung von Eigenwertproblemen in der Elastokinetik. By K. Hohenemser. [Pp. 89, with 15 figures.] (Berlin: Ergebnisse der Mathematik. Springer, 1932. RM. 10.50.)

THE classical study of the oscillations of elastic systems like strings, rods, membranes, plates, bowls, etc., by Helmholtz, Kirchhoff, Rayleigh, and the generalized mathematical treatment of oscillations in continuous elastic media due to Hilbert, Fredholm, etc., are insufficient for the treatment of the problems involving oscillations that arise in modern technology, where the conditions are not as ideal as in the classical problems, and where, further, the engineer needs comparatively rough but quick methods of dealing with all cases that occur in practice. Dr. Hohenemser's aim

is to give an account of the general approximate methods available for such problems, so that the mathematician may see the practical applications of the mathematical theory, and the engineer may become acquainted with the fundamental methods.

The author therefore gives a brief account of the relevant material out of the theory of linear integral equations of the second kind, although he utters a warning against over-estimating the practical value of the beautiful and compact mathematical methods, and advises the use of the differential equations and the boundary conditions in most cases. He gives in some detail an account of the recent work on the iterative methods for approximating to the fundamental period, due to Picard, Ritz, Mises, van den Dungen, Krilloff, Courant, Trefftz, etc., and then considers the method of replacing an actual problem by one very much like it, but simplified in such a way as to make the mathematical treatment easier—*e. g.*, considering the elastic material to be concentrated in a finite number of point masses; or replacing variable line density and variable rigidity by uniform line density and rigidity; or defining a system whose fundamental period is nearly the same as that of an assigned overtone of the given system; or considering a given oscillating system as the combination of simpler systems.

Practical applications occupy the remainder of the book. Brief, often rather summary, accounts are given of a number of problems to which the approximate methods have been applied, mainly in connexion with revolving parts of machinery, oscillations in foundations of buildings containing heavy machinery, turbine driven steamers, etc. In view of the widely spread literature in a branch of engineering which affords almost unlimited scope for the application of the mathematical processes, one must express gratitude to Dr. Hohenemser for his clear account of an intricate subject.

Incidentally one cannot help wondering whether sufficient has been done in this country in order to bring mathematics and engineering together. Here is a vast field of research of great importance to technology and industry where modern mathematical methods play an important rôle, to the mutual advantage of the engineer and the mathematician.

Lamésche, Mathieusche und verwandte Funktionen in Physik und Technik. By M. J. O. Strutt. [Pp. 116, with 12 figures.] (Berlin: Ergebnisse der Mathematik, Springer, 1932. RM. 13.60.)

Although only six or seven years have elapsed since connected accounts of Lamé and of Mathieu functions appeared in Whittaker and Watson's 'Modern Analysis,' and in

Humbert's 'Fonctions de Lamé et fonctions de Mathieu,' mathematicians and physicists will be grateful to Dr. Strutt for this exhaustive study of the subject from the point of view of applications to mechanical, physical, and technical problems. Having shown how Lamé's and Mathieu's differential equations arise out of a variety of problems in wave propagation and oscillations in spaces bounded by ellipsoids or elliptic cylinders, in hydrodynamical problems with analogous boundaries, in wave mechanics, etc., the author discusses the famous differential equation introduced by the American, G. W. Hill, half a century ago in order to deal with the motion of the moon. This differential equation is equivalent to the statement that the ratio of the second differential coefficient of a function to the function itself is a periodic function of the independent variable, and includes, as special cases, the equations defining the Legendre polynomials and associated functions, and also the equations of Bessel, Schrödinger, etc. The methods deduced in connexion with Hill's equation and the well-known infinite determinant, which, when equated to zero, yields the characteristic exponent of the solution in terms of series of exponential functions, are then applied to the study of Mathieu's equation, in which the periodic function of the independent variable consists of two terms, one a constant, and the other a constant times a circular function of the independent variable.

Arithmetical computation is an important feature of the application of Mathieu functions to problems in applied mathematics and physics, and considerable evolution in this respect has taken place in the last few years. Mathieu's calculations of sixty years ago have been adapted, particularly through the work of Whittaker, Ince, and Goldstein—all in this country,—so as to apply to all values of the constants in the equation. The importance of Ince's and Goldstein's work lies in the actual arithmetical application of their methods, since the mathematical basis of their methods was given in the classical treatise of Heine last century.

Lamé's differential equation can be roughly defined as Hill's equation, in which the second differential coefficient of a function bears to the function itself a ratio represented by a doubly periodic function of the independent variable. It occurs, for example, in the study of Laplace's equation of the potential function associated with an ellipsoidal boundary, and a special case is when the ellipsoid is a spheroid. The representation of Lamé's functions in terms of spherical harmonics, as deduced by C. Niven and R. Maclaurin, and in terms of Bessel functions or Hankel functions, is discussed, and the relation to Mathieu functions as degenerate forms of Lamé functions demonstrated.

Applications to a number of physical and other problems are indicated, such as diffraction of electromagnetic or sound-waves through an elliptical aperture in a flat screen, or round an ellipsoid or an elliptic cylinder; propagation of sound produced by a symmetrically vibrating disk, or through a "horn" defined by a hyperboloid of one sheet; oscillations in air bounded by an ellipsoid; induction in an ellipsoidal conductor; motion of an electron in a one-dimensional electric field, and the theory of the "wave-sieve" due to the author himself; and, finally, the quantum mechanics of the gyroscope as developed by Kramers and Ittmann.

Dr. Strutt pays more attention to the mechanical and physical applications than to the pure mathematical theory. The work is a report rather than a treatise on the subject. As such it is a useful addition to the relatively older literature on the subject, and a valuable collation of recent work in its field.

Hearing in Man and Animals. By R. T. BEATTY, M.A., B.E., D.Sc. [Pp. 227+99 figs.] (G. Bell & Sons. Price 12*s.* net.)

"THE purpose of this book is to convey in plain terms to the general reader a connected account of the phenomena of audition in living creatures, and of the various mechanisms by which animals are made sensitive to the range of sounds which is important for their welfare. As the author progressed . . . in collecting and correlating data supplied by the researches of anatomists, physiologists, physicists, engineers and psychologists he could not fail to note a lack of interchange of ideas between these specialized classes of workers, a lack which seems to be due partly to intense absorption in limited fields of investigation, but in greater degree to difficulty in comprehending the formidable technical terms peculiar to each subject."

So says the author in his preface to this admirable book.

In collecting the data rendered so easily available in the pages of this well-written volume, Dr. Beatty has done a real service to those whom he undoubtedly had in mind—the general scientific reader.

The author's treatment of the subject matter is logical, of course, and starts with the ear in human beings regarded mainly from the anatomical and evolutionary standpoints. Theories of hearing are then examined in detail. A study of the sensations of hearing follows, a chapter which is particularly interesting in its wide considerations. Thresholds of feeling and hearing; the bel as a unit; resonance; beats; combination tones; trills; the ear as possessing the faults of a badly designed loud-speaker; sense of direction; mental

effects ; fatigue ; these are some of the topics discussed in this wide-flung chapter. Successive sections deal with hearing in animals, music, noise, and defects in hearing.

The writer is a humorist as well as a professional physicist, and has skilfully amused himself—and his readers for sure—by an unusual, but delightful, blend of humour and science.

His " ladder of noise " is a case in point—and he takes an obvious delight in pointing out that whereas " the loudest and most awe-inspiring sound in Nature is the lion's roar " (it would be churlish of the reader even to think of Niagara !), yet some motor horns are of thirty lion power, a noise which comes perilously close to the top of his ladder.

Dr. Beatty is clearly very concerned about the noises of civilisation, and he has much of interest to say about this important matter.

A fascinating and instructive description of the nature and evolution of hearing, this book can be confidently recommended to all interested in the subject.

The Crystalline State. Edited by Sir W. H. Bragg and W. L. Bragg.—Vol. I. *A General Survey*, by W. L. Bragg. [Pp. xiv.+352.] (G. Bell & Sons, Ltd. Price 26*s.*)

This book is the first of three volumes edited by Sir W. H. Bragg and W. L. Bragg which, when complete, will give a comprehensive account of our present knowledge of the crystalline state as derived by the methods of X-ray and electron analysis. During the twenty years which have elapsed since their earlier book, ' X-rays and Crystal Analysis,' was first published, the subject has developed enormously and become of interest to workers in almost every branch of science ; physicists, chemists, metallurgists, engineers, and, of late, even biologists have utilized the methods of X-ray analysis. In view of the wide general interest of the subject on the one hand, and of the highly specialized nature of many aspects of the subject on the other hand, the problem of bringing the earlier book up to date must have been a difficult one. But it has been solved in a novel and—so far as can be judged from the first volume—a highly successful manner. The first volume is complete in itself and gives an outline of the methods of crystal analysis, a summary of the more important results, and a description of some of the applications of these methods in other fields of investigation ; it might be described as a book within a book. The two later volumes will contain sections dealing in more detail with particular branches of the subject, such as the theory of space groups, the technique of crystal analysis, and the results of crystal analysis. Students and others requiring a general account of the subject will be interested mainly in the first volume ; research workers and

those requiring more detailed information will probably be more interested in the later volumes. Although this arrangement will entail a certain amount of over-lapping, the general plan is excellent, for it is the only way by which the general reader and the more specialized reader can both be adequately served; it is a method which might well be adopted by other writers whose subjects are of both general and specialized interest.

The first six chapters of the present volume deal with the lattice theory of crystals, the diffraction of X-rays by crystals, and the principal methods of analysis. The seventh chapter summarizes the results of crystal analysis. This is just half of the book, and it brings up to date the matter given in the earlier book, 'X-rays and Crystal Structure.' The latter half of the book has five chapters describing more recent developments, which will be specially interesting to many readers. There is a chapter on "Crystal Texture" which describes the application of X-ray methods to the investigation of substances such as glass, rubber, wool, silk, etc., which are not truly crystalline, but which have a sufficiently ordered arrangement of molecules to give diffraction effects. The same chapter also shows how X-ray methods have yielded valuable information regarding the large-scale properties of metals. Another chapter deals with X-ray optics, and especially with the recent advances made by Prof. W. L. Bragg, who has applied the well-known Abbé theory of the microscope to the corresponding X-ray problem. The diffraction of electrons by crystals is discussed in another chapter. Finally, in the form of appendices, data on the production of X-rays, X-ray wave-lengths, absorption coefficients, atomic scattering factors, etc., are given.

The book is very beautifully produced and is copiously illustrated with diagrams and photographs. The latter call for special comment, for they are all very fine examples of X-ray and electron photographs, and they have been magnificently reproduced. Except in Ch. XII., which gives an historical survey of the subject, references have not been given; this is unfortunate, for even though the book is intended primarily for the general reader, such a reader may want to consult some of the original papers. The text is singularly free from errors; the name of one of the authors of the diagram on p. 220 has, however, been omitted. Finally, it is only necessary to add that the book is a worthy successor to 'X-rays and Crystal Structure,' and to express the hope that we shall not have long to wait for the two succeeding volumes.

[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]